

```

DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF05383; la; 1.
DR Pfam: PF00076; rim; 1.
DR PRINTS: PR00302; LUPUSLA.
DR SMART: SM00715; LA; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS50102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; 1.
DR PROSITE: PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 206 AA; 23992 MW; 965B62F7DFPB90E9 CRC64;

Query Match
Best Local Similarity 57.1%; Score 47; DB 13; Length 206;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 ERAIROVEYFGD 15
DB 14 ESKICQIEYFGN 27

RESULT 14
QEMTR3
AC Q8MYR3; PRELIMINARY; PRT; 269 AA.
ID Q8MYR3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE S018087P (Fragment).
GN LARP OR CG14066.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dreesnek D., Farfan D., Friese E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celinker S.;
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY113653; AAM29658.1; -.
DR Flybase: FBgn0040108; larp.
DR InterPro: IPR006630; Larp.
DR Pfam: PF05383; la; 1.
DR SMART: SM00715; LA; 1.
FT NON TER 269
SQ SEQUENCE 269 AA; 29223 MW; 5295D9EABF04922 CRC64;

Query Match
Best Local Similarity 56.0%; Score 47; DB 5; Length 269;
Matches 11; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

QY 2 ERAIROVEYF-----GDF 16
DB 22 KEAIKQVEYFSDNLTGDF 42

RESULT 15
Q7XUK8
ID Q7XUK8 PRELIMINARY; PRT; 435 AA.
AC Q7XUK8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE OSUNBA0067K08.21 protein.
GN OSUNBA0067K08.21.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_Taxid=4530;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RA Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,
RA Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y.,
RA Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W.,
RA Sheng H.H., Gu J.L., Chen S.T., Ni L., Zhu F.H., Han B., Feng Q.,
RA Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X., Liu Y.L., Mu J., Yu Z.,
RA Chen L., Fan D.L., Meng Q.J., Zhang L., Lu Y.Q., Yu S.L., Liu X.H.,
RA Lu T.T., Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,
RA Qian Y.M., Ying K., Zhou B., Chen Z.H., Hao P., Zhang L., Wu M.,
RA Zhang R.Q., Guan J.P., Hong G.F.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL606627; CAD41244.1; -.
SQ SEQUENCE 435 AA; 48295 MW; 66BA0B471B0AE119 CRC64;

Query Match
Best Local Similarity 56.0%; Score 47; DB 10; Length 435;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 RAIROVEYFGD 15
DB 17 KEVLRQVEYFSD 29

```

Search completed: September 10, 2004, 18:00:18
 Job time : 37.648 secs

SQ	SEQUENCE	433 AA;	48095 MW;	BSE6EBAF51C35A8F7	CRC64;
OY	Query Match	57.1%;	Score 48;	DB 10;	Length 433;
ID	Best Local Similarity	53.8%;	Pred. No. 8.9;		
AC	Matches	7;	Conservative	4;	Mismatches 2; Indels 0; Gaps 0;
Db	12 KTVLROVEFYPSD	24			
RESULT 11					
ID	08L7B4	PRELIMINARY;	PRT;	433 AA.	
AC	08L7B4				
DT	01-OCT-2002 (TREMBLrel. 22, Created)				
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)				
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)				
DE	Hypothetical protein.				
GK	AAG32720.				
OC	Arabidopsis thaliana (Mouse-ear cress).				
CC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
CC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
CC	eucosids II; Brassicales; Brassicaceae; Arabidopsids.				
OX	NCBI_TaxID=3702;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,				
RA	Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,				
RA	Deng J.M., Hayashizaki Y., Huan V.W., Lee J.M., Ishida J., Kamiya A.,				
RA	Kawai J., Kim C.J., Narusaka M., Quach H.L., Sakurai T., Satou M.,				
RA	Seki M., Shim P., Tang C.C., Toriumi M., Wallender E.K., Wong C.,				
RA	Wu H.C., Yamada K., Yu G., Yuan S., Shinozaki K., Eckert J.,				
RA	Theologis A., Davis R.W.;				
RL	submitted (JUL-2002) to the EMBL/genbank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,				
RA	Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,				
RA	Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,				
RA	Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,				
RA	Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,				
RA	Eckert J., Theologis A., Davis R.W.;				
RL	Submitted (SEP-2002) to the EMBL/genbank/DBJ databases.				
DR	EMBL; AY136302; AAM96968.1; -				
DR	EMBL; BT000396; AAN15715.1; -				
DR	GO; GO:0005634; C:nucleus; IEA.				
DR	GO; GO:0003723; F:RNA binding; IEA.				
DR	InterPro; IPRO02344; Lupus Ia.				
DR	InterPro; IPRO06630; Lupus_Ia_dom.				
DR	InterPro; IPRO00504; RNA_rec_mot.				
DR	pfam; PF05383; La; 1.				
DR	pfam; PF00076; rrm; 1.				
DR	PRINTS; PR00302; LUPUSLA.				
DR	SMART; SM00715; LA; 1.				
DR	SMART; SM00360; RRM; 1.				
DR	PROSITE; PS50102; RRM; 1.				
KW	Hypothetical protein.				
SQ	SEQUENCE	433 AA;	48126 MW;	CFF6F11A29AA0318	CRC64;
Query Match					
		57.1%;	Score 48;	DB 10;	Length 433;
Best Local Similarity		53.8%;	Pred. No. 8.9;		
Matches		7;	Conservative	4;	Mismatches 2; Indels 0; Gaps 0;
OY	3 RAIIROVEYFGD	15			
ID	:::				
Db	12 KTVLROVEYFGD	24			
RESULT 12					
ID	065529	PRELIMINARY;	PRT;	483 AA.	
AC	065529;				

```

DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN FAD11.80 OR AY4932720.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Beyer M., Beyer V., Rechmann S., Borkova D., Ansgore W., Hoheisel J.,
RA Mewes H.W., Mayer K.F.X., Scheller C.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Beyer V., Rechmann S., Borkova D., Ansgore W., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL022537; CAI18589.1; -.
DR EMBL; AL161582; CAB79989.1; -.
DR PIR; T04453; T04453.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR Pfam; PF00076; Irm; 1.
DR PRINTS; PRO0302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
KW Hypothetical protein.
SQ
SEQUENCE 483 AA; 54125 MW; 20840B34A0BC94E2 CRC64;

Query Match 57.1%; Score 48; DB 10; Length 483;
Best Local Similarity 53.8%; Pred. No. 9.9;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 RAIRQVEYFPGD 15
DB 12 KTVLRQVEYFPGD 24

RESULT 13
Q8QHI5 PRELIMINARY; PRT; 206 AA.
AC Q8QHI5;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Untranslated region binding-protein.
DE UBP.
GN UBP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA L'Ecuquer T.J., Fang H.-L.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF467897; AAL76269.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.

```


Best Local Similarity 57.1%; Pred. No. 1.1;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 ERAIRQVEYFFGD 15

Db 11 EKKVABQIEYFFGD 24

RESULT 5

Q9FL36 PRELIMINARY; PRT; 411 AA.

AC Q9FL36; (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Similarity to RNA-binding protein.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.

NCBI_TaxID=3702;

QY SEQUENCE FROM N.A.

RP STRAIN=Columbia;

RC MEDLINE=98344145; PubMed=9679202;

RX Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,

Tabata S.;

RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence

features of the regions of 1,381,565 bp covered by twenty one

physically assigned P1 and TAC clones."

RL DNA Res. 5:131-145(1998).

DR EMBL: AB010698; BAB1080.1; -

DR GO: GO:0005634; C:nucleus; IEA.

DR GO: GO:0003723; F:RNA binding; IEA.

DR InterPro: IPR002344; Lupus_La.

DR InterPro: IPR006630; Lupus_La_dom.

DR InterPro: IPR005054; RNA_rec_mot.

DR Pfam: PF05383; La; 1.

DR Pfam: PF00076; rtm; 1.

DR PRINTS: PR00302; LUPUSLA.

DR SMART: SM00715; LA; 1.

DR SMART: SM00360; RRM; 1.

DR PROSITE: PSS0102; RRM; 1.

DR PROSITE: PSS0102; RRM; 1.

DR PROSITE: PSS0102; RRM; 1.

DR PROSITE: PSS0102; RRM; 1.

RA Hiranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
Pam P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;

RT "Arabidopsis cDNA clones."

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,

Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,

Huan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,

Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,

Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,

Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,

Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,

Ecker J.R.;

RT "Arabidopsis ORF clones."

RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY050403; AAK91419.1; -

DR EMBL: BT000588; AAK18157.1; -

DR GO: GO:0005634; C:nucleus; IEA.

DR GO: GO:0003723; F:RNA binding; IEA.

DR InterPro: IPR002344; Lupus_La.

DR InterPro: IPR006630; Lupus_La_dom.

DR InterPro: IPR005054; RNA_rec_mot.

DR Pfam: PF05383; La; 1.

DR Pfam: PF00076; rtm; 1.

DR PRINTS: PR00302; LUPUSLA.

DR SMART: SM00715; LA; 1.

DR SMART: SM00360; RRM; 1.

DR PROSITE: PSS0102; RRM; 1.

DR PROSITE: PSS0102; RRM; 1.

DR PROSITE: PSS0102; RRM; 1.

DR PROSITE: PSS0102; RRM; 1.

DR PROSITE: PSS0102; RRM; 1.

DR PROSITE: PSS0102; RRM; 1.

DR PROSITE: PSS0102; RRM; 1.

QY 5 IIRQVEYFFGD 15

Db 107 IIRQVEYFFGD 117

Query Match 63.1%; Score 53; DB 10; Length 422;

Best Local Similarity 90.9%; Pred. No. 1.2;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IIRQVEYFFGD 15

Db 107 IIRQVEYFFGD 117

Query Match 63.1%; Score 53; DB 10; Length 422;

Best Local Similarity 90.9%; Pred. No. 1.2;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IIRQVEYFFGD 15

Db 107 IIRQVEYFFGD 117

Query Match 63.1%; Score 53; DB 10; Length 422;

Best Local Similarity 90.9%; Pred. No. 1.2;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 IIRQVEYFFGD 15

Db 107 IIRQVEYFFGD 117

Query Match 63.1%; Score 53; DB 10; Length 422;

Best Local Similarity 90.9%; Pred. No. 1.2;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Query Match	Similarity	92.9%	Score 78:	DB 5:	Length 390:
Best Local	Similarity	100.0%	Pred. No. 5.9e-05:	Mismatches 0:	Indels 0:
Matches 15:	Conservative	0:	Mismatches 0:	Indels 0:	Gaps 0:
Db	50	QERAIRROVEYYFGD 64			
RESULT 2					
Q9CYB9	PRELIMINARY:	PRT: 381 AA.			
AC	Q9CYB9:				
DT	01-JUN-2001 (TREMBLrel. 17, Created)				
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)				
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)				
DE	Sjogren syndrome antigen B.				
GN	SSB				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
NCBI_TaxId=10090;					
RP	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-C57BL/6J; Tissue=Embryo;				
RA	MEDLINE=21085660; PubMed=11217851;				
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,				
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi S., Fukuda S.,				
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondou S., Yamanaoka I.,				
RA	Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,				
RA	Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,				
RA	Fleischmann W., Gasteierland T., Gissi C., King B., Kochava H.,				
RA	Kuehl P., Lewis S., Matsumoto Y., Nikaide I., Pesole G., Quackenbush J.,				
RA	Schmitt L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,				
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,				
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,				
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,				
RA	Guerninich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,				
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mombaerts P.,				
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,				
RA	Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,				
RA	Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,				
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kontsuki S.,				
RA	Hayashizaki Y.;				
RT	"Functional annotation of a full-length mouse cDNA collection";				
RL	Nature 409:685-690 (2001).				
EMBL	AK017823; BAB30957.1; -				
DR	MGI: 98423; Ssb				
DR	GO: 0005634; C:nucleus; IDA.				
DR	InterPro: IPR002344; Lupus_La.				
DR	InterPro: IPR006630; Lupus_La.dom.				
DR	InterPro: IPR000504; RNA_rec_mot.				
DR	Pfam: PF05383; La: 1.				
DR	Pfam: PF00076; rrm; 1.				
DR	PRINTS: PR00302; LUPUSLA.				
DR	SMART; SM00715; LA: 1.				
SEQ	SEQUENCE 381 AA; 43891 MW; 2E2DEF1452C0F0E9 CRC64;				
Query Match	67.9%	Score 57:	DB 11:	Length 381:	
Best Local Similarity	66.7%	Pred. No. 0.22:			
Matches 10:	Conservative 1:	Mismatches 4:	Indels 0:	Gaps 0:	
Db	14	ERAIRROVEYYFGD 16			
QY	2	ERAIRROVEYYFGD 16			
DB	14	BAKICHOIEYRGDF 28			
RESULT 3					
Q8BTU4	PRELIMINARY:	PRT: 415 AA.			
AC	Q8BTU4:				
DT	01-MAR-2003 (TREMBLrel. 23, Created)				
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)				

```

DE 01-OCR-2003 (TREMBLrel. 25, last annotation update)
DE Sjogren syndrome antigen B.
CN SSB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium,
RA the RIKEN Genome Expression Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RL EMBL; AK088677; BAC04098.1; -.
DR MGD; MGI:98423; Ssd.
DR GO; GO:0005634; C:nucleus; IDA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF05383; Ia; 1.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 415 AA; 47657 MW; A7545C7686AC8363 CRC64;

Query Match: 67.9%; Score 57; DB 11; Length 415;
Best Local Similarity 66.7%; Pred. No. 0.24;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Cy 2 ERATIRQVEYYFGDF 16
Db 14 EAKIHOIEYYGDF 28

RESULT 4
07ZTIO PRELIMINARY; PRT; 401 AA.
ID 07ZTIO.
AC 07ZTIO.
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCR-2003 (TREMBLrel. 25, Last annotation update)
DE Similar to sjogren syndrome antigen B (Autoantigen Ia) .
OS Brachydanio retio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RL Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC045392; AAH45392.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 401 AA; 46138 MW; 2BF032FDD3916291 CRC64;

Query Match: 63.1%; Score 53; DB 13; Length 401;

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:29:01 / Search time 36.648 Seconds

(without alignments)
137.751 Million cell updates/sec

Title: US-09-836-073-19

Sequence: 1 QERAIRQVEYRFGDF 16

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Listing first 45 summaries

SPREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	92.9	390	5	Q8T8V5
2	57	67.9	381	11	Q9CYB9
3	57	67.9	415	11	Q8BTU4
4	53	63.1	401	13	Q7ZTU0
5	53	63.1	411	10	Q9FL36
6	53	63.1	422	10	Q94A38
7	50	59.5	396	5	001806
8	49	58.3	389	10	Q8S0T8
9	48	57.1	427	13	Q7ZTK2
10	48	57.1	433	10	Q93ZV7
11	48	57.1	433	10	Q8L7B4
12	47	56.0	206	13	Q8QHI5
13	47	56.0	269	5	Q8MYR3
14	47	56.0	435	10	Q7XU88
15	47	56.0	545	10	Q80567
16	47	56.0	545	10	Q80567

17	47	56.0	826	10	Q940X9	Q940X9 arabidopsis
18	47	56.0	965	5	Q9VAM5	Q9VAM5 drosophila
19	47	56.0	1403	5	Q9NH66	Q9NH66 drosophila
20	47	56.0	1403	5	Q8IMM4	Q8IMM4 drosophila
21	47	56.0	1442	5	Q86B76	Q86B76 drosophila
22	45.5	54.2	642	16	Q9KEU7	Q9KEU7 bacillus ba
23	45	53.6	491	10	Q8LMP9	Q8LMP9 oryza sativ
24	45	53.6	914	12	Q85427	Q85427 rat cytochrome
25	43	51.2	343	2	Q9XB17	Q9XB17 bacillus ce
26	43	51.2	373	16	Q8CU33	Q8CU33 staphylococ
27	43	51.2	469	16	Q66887	Q66887 aquifex aeo
28	43	51.2	669	16	Q8XRQ1	Q8XRQ1 ralatonia s
29	43	51.2	833	16	Q8LUP0	Q8LUP0 bacillus an
30	43	51.2	846	16	Q9RT05	Q9RT05 delnocooc
31	42.5	50.6	175	5	Q8I979	Q8I979 plasmodium
32	42	50.0	166	16	Q87NT5	Q87NT5 vibrio para
33	42	50.0	161	16	Q82279	Q82279 enterococu
34	42	50.0	213	12	Q9YFX1	Q9YFX1 foot-and-mo
35	42	50.0	236	11	Q9CTN3	Q9CTN3 mus musculu
36	42	50.0	343	10	Q94LD0	Q94LD0 oryza sativ
37	42	50.0	395	5	Q8I9V8	Q8I9V8 manduca sex
38	42	50.0	398	2	Q7X4V8	Q7X4V8 escherichia
39	42	50.0	398	16	Q7UAP6	Q7UAP6 shigella fl
40	42	50.0	407	16	Q83IM4	Q83IM4 mus musculu
41	42	50.0	492	11	Q9D3J0	Q9D3J0 mus musculu
42	42	50.0	492	11	Q8CA51	Q8CA51 mus musculu
43	42	50.0	492	11	Q8C9A3	Q8C9A3 mus musculu
44	42	50.0	492	11	Q8BN59	Q8BN59 mus musculu
45	42	50.0	788	17	Q58603	Q58603 pyrococcus

ALIGNMENTS

RESULT 1

Q8T8V5 PRELIMINARY; PRT; 390 AA.

AC Q8T8V5;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE AT22034D.
GN LA OR CG10922.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Streptococcus M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Friese E.,
RA George R., Gonzalez W., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Cejnkner S.,
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

EMBL; AY075257; AL68124.1; -
DR FLYBase; FBgn0011638; La.
DR GO; GO:0008098; F:5S rRNA primary transcript binding; IDA.
DR GO; GO:0003723; F:RNA binding; NAS.
DR InterPro; IPR002344; Lupaue La.
DR InterPro; IPR006630; Lupaue La dom.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR Pfam; PF00076; rtm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS01021; RRM; 1.
DR PROSITE; PS00303; RRM_RNP_1; 1.
SQ SEQUENCE 390 AA; 44869 MW; 797FDE26B903C909 CRC64;

```

RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
RA Apodaca J., Mancharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blatter F.R.,
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RA Nature 409:529-533 (2001).
RL
RN
RP
RC STRAIN=O157:H7 / RIMD 0509952;
RC MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shingawa H.;
RA "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RT DNA Res. 8:11-22 (2001).
CC
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE005657; AAG59456.1; ALT_INTT.
DR EMBL; AE002569; BAB38657.1; -.
DR PIR; B91283; B91283.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 25 45
FT TRANSMEM 74 94
FT TRANSMEM 96 116
FT TRANSMEM 143 163
FT TRANSMEM 176 196
FT TRANSMEM 229 249
FT TRANSMEM 279 299
FT TRANSMEM 334 354
SQ SEQUENCE 398 AA; 44817 MW; C43710A463D289C9 CRC64;

Query Match 50.0%; Score 42; DB 1; Length 398;
Best Local Similarity 57.1%; Pred. No. 8.5;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 QERRAIRQVEYYFG 14
DB 276 QRRMTIAQLIYVFG 289

RESULT 15
YJGN_ECOLI STANDARD; PRT; 398 AA.
ID YJGN_ECOLI AC P39338; P39339; P76811;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yjgn.
GN YJGN OR B4257.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN
RP
RC STRAIN=K12 / MG1655;
RC MEDLINE=85334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blatter F.R.,
RA "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes.";
RT Nucleic Acids Res. 23:2105-2119 (1995).
RN
RP REVISIONS.

```

```

RC STRAIN=K12 / MG1655;
RC MEDLINE=97426617; PubMed=9278503;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kiropatrick H.A., Goeden W.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12.";
RT Science 277:1453-1474 (1997).
RL
RN
RP
RC CONCEPTUAL TRANSLATION.
RC Rudd K.E.;
RL Unpublished observations (JAN-2000).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT HAD TO
CC BE INTRODUCED IN POSITION 12 TO PRODUCE THIS ORF.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U14003; AAA97153.1; ALT_FRAME.
DR EMBL; U14003; AAA97154.1; ALT_FRAME.
DR Ecogene; EG12533; YJGN.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 25 45
FT TRANSMEM 74 94
FT TRANSMEM 96 116
FT TRANSMEM 143 163
FT TRANSMEM 176 196
FT TRANSMEM 229 249
FT TRANSMEM 279 299
FT TRANSMEM 334 354
SQ SEQUENCE 398 AA; 44877 MW; 814BFAD98E816A63 CRC64;

Query Match 50.0%; Score 42; DB 1; Length 398;
Best Local Similarity 57.1%; Pred. No. 8.5;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 QERRAIRQVEYYFG 14
DB 276 QRRMTIAQLIYVFG 289

Search completed: September 10, 2004, 17:53:13
Job time : 7.43575 secs

```

LAH1 YEAST
 ID LAH1 YEAST STANDARD; PRT; 275 AA.
 AC P33399;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE LAH1 OR LHPI OR YLAI OR YDL051W.
 GN LAH1 OR LHPI OR YLAI OR YDL051W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Bacteria; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=YPH501;
 RX MEDLINE=94309661; PubMed=8035818;
 RA Yoo C.J., Molin S.L.;
 RT "La proteins from *Drosophila melanogaster* and *Saccharomyces cerevisiae*: a yeast homolog of the La autoantigen is dispensable for growth."
 RL Mol. Cell. Biol. 14:5412-5424(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1678;
 RX MEDLINE=97313263; PubMed=9169867;
 RA Jacq C., Alt-Moerbe J., Andre B., Arnold W., Bahr A., Ballesta J.P.G., Barques M., Baron L., Becker A., Bibeau N., Bloeker H., Blugeon C., Boskovic J., Brandt P., Bruckner M., Buitrago M.J., Coster F., Delaveau T., del Rey F., Dufon B., Eide L.G., Garcia-Cantalejo J.M., Goffeau A., Gomez-Paris A., Granotier C., Hanemann V., Hankeln T., Hoebiel J.D., Jaeger W., Jimenez A., Jomiaux J.-L., Kraemer C., Kuester H., Laananen P., Legros Y., Louis E.J., Moeller-Rieker S., Monnet A., Moro M., Mueller-Auer S., Nussbaumer B., Paricio N., Paulin L., Pera J., Perez-Alonso M., Perez-Ortin J.E., Pohl T.M., Prydz H., Purnelle B., Rasmussen S.W., Remacha M., Revuelta J.L., Rieger M., Salom D., Saluz H.P., Salz J.E., Saren A.-M., Schaefer M., Scharte M., Schmidt E.R., Schneider C., Scholler P., Schwarz S., Urrestarazu L.A., Verhasselt P., Vissers S., Voet M., Volckaert G., Wagner G., Wambuit R., Wedler E., Wedler H., Woelfl S., Harris D.E., Bowman S., Brown D., Churcher C.M., Connor R., Dedman K., Gentles S., Hamlin N., Hunt S., Jones L., McDonald S., Murphy L., Niblett D., Odell C., Oliver K., Rajandream M.A., Richards C., Shore L., Walsh S.V., Barrall B.G., Dietrich F.S., Milligan J.T., Allen E., Aranjó R., Aviles E., Bero A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M., Hunnicke-Smith S., Hyman R.W., Komp C., Laekkeri D., Lew H., Lin D., Mosedale D., Nakahara K., Namach A., Oehner P., Oh C., Petel F.X., Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N., Winant A., Yelton M.A., Botstein D., Davis R.W., Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Favello A., Fulton L., Gattung S., Greco T., Hallsworth K., Hawkins J., Hillier L.W., Jier M., Johnson D., Johnston L., Kirsten J., Kucaba T., Langston Y., Latreille P., Le T., Mardis E., Meneses S., Miller N., Nhan M., Pauley A., Peluso D., Rifkin L., Riles L., Tach A., Trevasis E., Vignati D., Wilcox L., Wohlman P., Vaudin M., Wilson R., Waterston R., Albertmann K., Hahl J., Heumann K., Kleine K., Mewes H.-W., Zollner A., Zaccaria P.;
 RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome IV."
 RL Nature 387:75-78(1997).
 RN [3]
 RP SEQUENCE OF 101-275 FROM N.A., AND CHARACTERIZATION.
 RX MEDLINE=95097387; PubMed=7799435;
 RA Lin-Marg N., Clarkson S.G.;
 RT "A yeast RNA binding protein that resembles the human autoantigen La."
 RL J. Mol. Biol. 245:81-85(1995).
 RN [4]
 RP SEQUENCE OF 1-252 FROM N.A.
 RX MEDLINE=94012814; PubMed=8408076;
 RA Naelek M.W., Wells G.B., Lester R.L., Dickson R.C.;
 RT "A suppressor gene that enables *Saccharomyces cerevisiae* to grow without making sphingolipids encodes a protein that resembles an *Escherichia coli* fatty acyltransferase."
 RL J. Biol. Chem. 268:22156-22163(1993).

RN [5]
 RP SIMILARITY TO PROTEIN LA.
 RA Baum B.;
 RL Unpublished observations (NOV-1993).
 CC -1- FUNCTION: Binds to the precursors of polymerase III RNAs. Binds preferentially to precursors ending in U residues.
 CC -1- SUBCELLULAR LOCATION: Nuclear (probable).
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcement/> or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; L33023; AAA21777.1; -
 DR EMBL; Z74099; CAA98612.1; -
 DR EMBL; L13282; AAA16515.1; -
 DR EMBL; X80801; CAA56782.1; -
 DR PIR; B48600; B48600.
 DR GERMOnline; 140293; -
 DR SGD; S0002209; LHP1.
 DR GO; GO:0005730; C:nucleolus; IDA.
 DR GO; GO:0005654; C:nucleoplasm; IDA.
 DR GO; GO:0003723; F:RNA binding; IDA.
 DR GO; GO:0008033; P:RNA processing; IMP.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
 KM RNA-binding; Nuclear Protein.
 FT DOMAIN 123 216 RNA-BINDING (RRM).
 FT FT 240 256 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 SQ SEQUENCE 275 AA; 32104 MW; 43CB2EC740978D CRC64;
 Query Match 50.0%; Score 42; DB 1; Length 275;
 Best Local Similarity 54.5%; Pred. No. 5.8;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 Oy 6 IRQVEYFYCDF 16
 Db 34 LKQVEFYFSEF 44
 ID YUGN ECOS7
 AC P58213; STANDARD; PRT; 398 AA.
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein YUGN.
 GN YUGN OR Y5869 OR EGS5234.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glaesner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: X59720; CAA42379.1; -
 CC PIR: S19365; S19365.
 CC GerMOnline: 138873; -
 CC SGD: S0000542; SRO9.
 CC DR GO: GO:0005844; C:polysome; IDA.
 CC DR GO: GO:0003723; F:RNA binding; IDA.
 CC DR GO: GO:0006412; P:protein biosynthesis; IPT.
 CC DR InterPro: IPR006630; Lupus_la_dom.
 CC DR Pfam: PF05383; La; 1.
 CC DR SMART: SM00715; La; 1.
 CC FT DOMAIN 200 203 POLY-GUN.
 CC FT 254 258 POLY-ASN.
 CC SQ SEQUENCE 466 AA; 51789 MW; AD8B089545117C CRC64;

Query Match 51.2%; Score 43; DB 1; Length 466;
 Best Local Similarity 63.6%; Pred. No. 6.7;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 5 IIRQVEYFGD 15
 Db 297 IARQIEYFSE 307

RESULT 12
 ID VP45_ARATH STANDARD; PRT; 569 AA.
 AC 049048; O80650;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 GN Vacuolar protein-sorting protein 45 homolog (AtVP45).
 OS VPS45 OR AT1G7140 OR T14N5.2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=98289086; PubMed=9625693;
 RA Bassham D.C., Raikhel N.V.;
 RT "An Arabidopsis VPS45p homolog implicated in protein transport to the
 RL vacuole.";
 RL Plant Physiol. 117:407-415 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gail J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
 RA Miliutecner J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Uterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
 RA Wu D., Yu G., Frazer C.M., Venter J.C., Davis R.W.,

RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana.";
 RT Nature 408:816-820 (2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=22954850; PubMed=14593172;
 RA Yamada K., Lim J., Dale J.M., Chen H., Shim P., Palm C.J.,
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
 RA Arakawa T., Banh J., Banno P., Bowser L., Brooks S.Y., Carninci P.,
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Guiral M., Hansen N.F.,
 RA Hayashizaki Y., Johnson-Hopson C., Hsu W.V., Iida K., Karnes M.,
 RA Khan S., Koesema B., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamita A., Meyers C., Nakajima M., Natusaka M., Seki M., Sakurai T.,
 RA Satou M., Tames R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Empirical analysis of transcriptional activity in the Arabidopsis
 RL genome.";
 RL Science 302:842-846 (2003).
 RN [4]
 RP INTERACTIONS WITH SYP21, SYP22, SYP31, SYP41, SYP42, SYP61, VTI11 AND
 RP VTI12.
 RX MEDLINE=20346955; PubMed=1088666;
 RA Bassham D.C., Sanderfoot A.A., Kovaleva V., Zheng H., Raikhel N.V.;
 RT "AtVPS45 complex formation on the trans-Golgi network.";
 RL Mol. Biol. Cell 11:2251-2265 (2000).
 CC -1- FUNCTION: Involved in the protein transport to the vacuole,
 CC probably at the level of vesicle fusion at the prevacuolar
 CC (TGN) and not in transport from the TGN to the prevacuolar
 CC compartment. Binds syntaxins.
 CC -1- SUBUNIT: Interacts with both SYP41 or SYP42 and VTI12, but in
 CC different domains of the trans-Golgi network. Does not interact on
 CC the prevacuolar compartment with VTI11, SYP21 or SYP22, or on the
 CC cis-Golgi with SYP1.
 CC -1- SUBCELLULAR LOCATION: Peripheral membrane protein. Binds to trans-
 CC Golgi network membranes through interaction with other proteins.
 CC -1- TISSUE SPECIFICITY: Highly expressed in roots, lower expression in
 CC leaves, stems and flowers.
 CC -1- SIMILARITY: Belongs to the STXBP/UNC-18/SEC1 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: AF036234; AAC39472.1; -
 CC DR EMBL: AC004260; AAC34344.1; -
 CC DR EMBL: AY050370; AAK91386.1; -
 CC DR EMBL: AY101517; AAM26638.1; -
 CC DR PIR: T00445; T00445.
 CC DR PIR: T52056; T52056.
 CC DR InterPro: IPR001619; Sec1-like.
 CC DR Pfam: PF00995; Sec1; 1.
 CC FT TRANSPORT: Protein transport; Golgi stack; Multigene family.
 CC FT CONFLICT 362 362 T -> I (IN REF. 1).
 CC SQ SEQUENCE 569 AA; 64942 MW; 844F24C9A21E9D1D CRC64;

Query Match 51.2%; Score 43; DB 1; Length 569;
 Best Local Similarity 40.0%; Pred. No. 8.2;
 Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

OY 2 ERAITRQVEYFGDF 16
 Db 116 EDEWVOQVEYVADF 130

RESULT 13

```

DR PIR; T43542; T43542.
DR GenEB Sprobe; SPACS7A10.10c; -.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La.dom.
DR InterPro; IPR005054; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS0102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
DR RNA-binding; Nuclear protein.
FT DOMAIN 154 236
FT CONFLICT 188 188 M -> I (IN REF. 1 AND 2).
SQ SEQUENCE 298 AA; 34616 MW; 64E6AB99340B87F4 CRC64;

Query Match 54.8%; Score 46; DB 1; Length 298;
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 ERATIRQVEYFGD 15
Db 64 EAEVTKQVEFYFSD 77

RESULT 10
HO2_RAT STANDARD; PRT; 315 AA.
ID_HO2_RAT
AC P23711;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Heme oxygenase 2 (EC 1.14.99.3) (HO-2).
GN HMOX2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=90237051; PubMed=2185251;
RA Rosenberg M.O., Maines M.D.;
RT "Isolation, characterization, and expression in Escherichia coli of a
RT cDNA encoding rat heme oxygenase-2."
RL J. Biol. Chem. 265:7501-7506(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=liver;
RX MEDLINE=94156193; PubMed=8112599;
RA McCoubrey W.K., Jr., Maines M.D.;
RT "The structure, organization and differential expression of the gene
RT encoding rat heme oxygenase-2."
RL Gene 139:155-161(1994).
RN [3]
RP SEQUENCE OF 142-232 FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=liver, and testis;
RX MEDLINE=88139412; PubMed=3343248;
RA Cruse I., Maines M.D.;
RT "Evidence suggesting that the two forms of heme oxygenase are
RT products of different genes."
RL products of different genes."
RN [4]
RP BIOL. Chem. 263:3348-3353(1988).
RL
-1- FUNCTION: Heme oxygenase cleaves the heme ring at the alpha
mechene bridge to form biliverdin. Biliverdin is subsequently
converted to bilirubin by biliverdin reductase. Under
physiological conditions, the activity of heme oxygenase is
highest in the spleen, where senescent erythrocytes are
sequestered and destroyed.
-1- FUNCTION: Heme oxygenase 2 could be implicated in the production
of carbon monoxide in brain where it could act as a
neurotransmitter.
-1- CATALYTIC ACTIVITY: Heme + 3 AH(2) + O(2) = biliverdin + Fe(2+) +

```

```

CC CO + 3 A + 3 H(2)O.
CC -1- SUBCELLULAR LOCATION: Microsomal.
CC -1- TISSUE SPECIFICITY: Widely distributed in body with a high
CC concentration in the brain.
CC -1- INDUCTION: Heme oxygenase 2 activity is non-inducible.
CC -1- SIMILARITY: Belongs to the heme oxygenase family.
CC -1- SIMILARITY: Contains 2 heme regulatory motif (HRM) repeats.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J05405; AAA41340.1; -.
DR EMBL; U05013; AAA19130.1; -.
DR EMBL; M18918; AAA41347.1; -.
DR PIR; A35199; A35199.
DR HSBP; P06762; IDVG.
DR InterPro; IPR002051; Heme_oxygenase.
DR Pfam; PR01126; Heme_oxygenase.
DR PRINTS; PR00086; HAEMOXIGINASE.
DR PROSITE; PS00593; HEME_OXYGENASE; 1.
DR Heme; Oxidoreductase; Microsome; Multigene family; Repeat.
FT METAL 44 44 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT REPEAT 263 268 HRM 1.
FT REPEAT 280 285 HRM 2.
FT CONFLICT 142 146 ONEPE -> EFRNK (IN REF. 3).
FT CONFLICT 230 232 MQI -> TER (IN REF. 3).
SQ SEQUENCE 315 AA; 35762 MW; 981ADE01DE1AFCF CRC64;

Query Match 52.4%; Score 44; DB 1; Length 315;
Best Local Similarity 40.0%; Pred. No. 3;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 QERATIRQVEYFGD 15
Db 104 RKEALIKQMEYFGE 118

RESULT 11
SR09_YEAST STANDARD; PRT; 466 AA.
ID_SR09_YEAST
AC J25567;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SR09 protein.
GN SR09 OR YCL037C OR YCL37C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Delgado M., Esteban M., Navas L.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=98043396; PubMed=9383048;
RA Kagami M., Tob-E A., Matsui Y.;
RT "SR09, a multicopy suppressor of the bud growth defect in the
RT Saccharomyces cerevisiae RHO3-deficient cells, shows strong genetic
RT interactions with tropomyosin genes, suggesting its role in
RT organization of the actin cytoskeleton."
RL Genetics 147:1003-1016(1997).
RN
-1- FUNCTION: May overlap in function with tropomyosin and may be
involved in organization of actin filaments. Acts as a multicopy
suppressor of RHO3.
-1- SIMILARITY: Some, to yeast SUP1.

```



```

RESULT 8
LAA_XENIA STANDARD; PRT; 428 AA.
ID LAA_XENIA
AC P28048;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE La protein homolog A (La ribonucleoprotein A) (La autoantigen
homolog A).
GN LAA1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RX MEDLINE=93287095; PubMed=8510143;
RA Scherly D., Stutz F., Lin-Marg N., Clarkson S.G.;
RT "La proteins from Xenopus laevis. cDNA cloning and developmental
expression."
RL J. Mol. Biol. 231:196-204(1993).
CC -1- FUNCTION: La protein plays a role in the transcription of RNA
polymerase III. It is most probably a transcription termination
factor. Binds to the 3' terminus of virtually all nascent
polymerase III transcripts (by similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- DEVELOPMENTAL STAGE: Barely detectable in stage I/II oocytes,
accumulate in stage III/IV oocytes, then exhibit a roughly
constant steady state level in mature oocytes, eggs, and early
embryos.
CC -1- PTM: Phosphorylated (Probable).
CC -1- MISCELLANEOUS: There are two forms of La, Laa and Lab, in xenopus.
CC -1- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X68817; CAA48715.1; -.
CC PIR; S33818; S33818.
CC InterPro; IPR002344; Lupus_La.
CC InterPro; IPR006630; Lupus_La_dom.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF00383; La; 1.
CC Pfam; PF00076; rrm; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00715; La; 1.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS50102; RRM; 1.
CC DR PROSITE; PS00030; RRM_RNP_1; 1.
CC RNA-binding; Nuclear protein; Phosphorylation.
CC FT DOMAIN 111 203 RNA-BINDING (RRM).
CC FT DOMAIN 196 232 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC FT DOMAIN 316 332 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC SQ SEQUENCE 428 AA; 48864 MW; AEB3A38B7DE3EC3 CRC64;
Query Match 57.1%; Score 48; DB 1; Length 428;
Best Match Similarity 72.7%; Pred. No. 0.81;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 5 IIRVEYFGD 15
DB 17 ICEQIEYFGD 27

```

```

RESULT 9
LAA1_SCHPO STANDARD; PRT; 298 AA.
ID LAA1_SCHPO
AC P87058; 013362; Q10458;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
GN SLAI OR SPAC57A10.10C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98067398; PubMed=9404894;
RA van Horn D.J., Yoo C.J., Xue D., Shi H., Wolin S.L.;
RT "The La protein in Schizosaccharomyces pombe: a conserved yet
dispensable phosphoprotein that functions in tRNA maturation."
RL RNA 3:1434-1443(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Utsunomiya R.R.U.;
RT Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Godle A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holyoak S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsle K.,
RA James K., Jones L., Jones M., Leathers S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymopoulos B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Meisel D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garçon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Koreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Pictakhin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -1- FUNCTION: Binds to the precursors of polymerase III RNAs.
CC -1- Functions in tRNA maturation.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF022949; AAB82145.1; -.
CC DR EMBL; AB011371; BAA24981.1; -.
CC DR EMBL; Z94864; CAB08173.1; -.
CC PIR; T38937; T38937.

```

DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PSS0102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 KW RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 111 187 RNA-BINDING (RRM).
 SQ SEQUENCE 415 AA; 47777 MW; 033FD9CC1E475F98 CRC64;

Query Match 67.9%; Score 57; DB 1; Length 415;
 Best Local Similarity 66.7%; Pred. No. 0.021;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ERAIR0VEYYFGDF 16
 DB 14 EAKIC0IEYYFGDF 28

RESULT 6
 LA_AEDAL STANDARD; PRT; 383 AA.

ID LA_AEDAL
 AC Q26457;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 OS Aedes albopictus (Forest day mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preyigota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.
 OX NCBI_TaxID=7160;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RX MEDLINE=96135233; PubMed=8551578;
 RA Pardigon N., Straus J.H.;
 RT "Mosquito homolog of the La autoantigen binds to Sindbis virus RNA."
 RT J. Virol. 70:1173-1181(1996).
 CC -1- FUNCTION: May be involved in transcription termination by RNA
 polymerase III. Binds RNA and DNA. Binds to the 3' end of the
 minus strand of Sindbis virus RNA. This may be significant for
 CC Sindbis virus RNA replication.
 CC -1- SUBCELLULAR LOCATION: Nuclear. Primarily nuclear, but significant
 CC amounts are present in the cytoplasm.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; S80954; AAB35931.1; -;
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rtm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PSS0102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
 KW RNA-binding; Nuclear protein; DNA-binding.
 FT DOMAIN 141 228 RNA-BINDING (RRM).
 SQ SEQUENCE 383 AA; 44430 MW; 4B5CC8F21C40F452 CRC64;

Query Match 66.7%; Score 56; DB 1; Length 383;
 Best Local Similarity 71.4%; Pred. No. 0.028;
 Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ERAIR0VEYYFGD 15
 DB 14 EAKIC0IEYYFGD 28

DB 44 EASTIR0LEYFGD 57

RESULT 7
 LAB_XENLA STANDARD; PRT; 427 AA.

ID LAB_XENLA
 AC P28049;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein homolog B (La ribonucleoprotein B) (La autoantigen
 DE homolog B).
 GN LAB1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8335;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocyte;
 RX MEDLINE=93267095; PubMed=8510143;
 RA Scherly D., Stutz F., Lin-Marg N., Clarkson S.G.;
 RT "La proteins from Xenopus laevis. cDNA cloning and developmental
 RT expression.";
 RT J. Mol. Biol. 231:196-204(1993).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA
 polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' terminus of virtually all nascent
 CC polymerase III transcripts (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (probable).
 CC -1- DEVELOPMENTAL STAGE: Barely detectable in stage I/II oocytes,
 CC accumulate in stage III/IV oocytes, then exhibit a roughly
 CC constant steady state level in mature oocytes, eggs, and early
 CC embryos.
 CC -1- PTM: Phosphorylated (probable).
 CC -1- MISCELLANEOUS: There are two forms of La, LaA and LaB, in xenopus.
 CC -1- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; X68818; CAA48716.1; -;
 DR PIR; S33817;
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rtm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PSS0102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
 KW RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 110 202 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 SQ SEQUENCE 427 AA; 48995 MW; 45F3146F934A355 CRC64;

Query Match 57.1%; Score 48; DB 1; Length 427;
 Best Local Similarity 72.7%; Pred. No. 0.81;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IIR0VEYYFGD 15
 DB 16 ICE0IEYYFGD 26

LA_MOUSE STANDARD; PRT; 415 AA.
 ID_ LA_MOUSE
 AC P32067;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus Ia protein homolog (La ribonucleoprotein) (La autoantigen homolog)
 GN SSB OR SS-B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93203630; PubMed=8454877;
 RA Topfer F., Gordon T., McCluskey J.;
 RT "Characterization of the mouse autoantigen La (SS-B). Identification of conserved RNA-binding motifs, a putative ATP binding site and reactivity of recombinant protein with poly(U) and human autoantibodies.";
 RT J. Immunol. 150:3091-3100 (1993).
 [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=FVB/N; Tissue=Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F., Diatchenko L., Mausina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Donald M.F., Casavant P.L., Scheetz T.E., Brownstein M.U., Uesdi T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Millar S.J., Bosak S.A., McMan P.J., McKernan K.J., Abjerson R.D., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Skalska D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [3]
 RN SEQUENCE OF 1-11 FROM N.A.
 RP Groetz D., Bachmann M.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' terminus of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including crRNA and 4.5S, 5S, 7S, and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: Phosphorylated (By similarity).
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; L00993; AAA9415.1; -
 DR EMBL; BC003820; AA03820.1; -
 DR EMBL; Y07951; CAA69249.1; -
 DR MGI; 98423; Ssb.

DR GO; 0005634; C:nucleus; IDA.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; xrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 111 187
 SO SEQUENCE 415 AA; 47756 MW; 2D75197692FDC933 CRC64;
 Query Match 67.9%; Score 57; DB 1; Length 415;
 Best Local Similarity 66.7%; Pred. No. 0.021;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 2 EBAIRQVEYRGDF 16
 DB 14 EAKICHQIEYRGDF 28
 ID LA_RAT STANDARD; PRT; 415 AA.
 AC P38656;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lupus Ia protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 GN SSB OR SS-B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93244655; PubMed=7916708;
 RA Semel I., Troester H., Bartsch H., Schwemmler M., Igloi G.L., Bachmann M.;
 RT "Isolation of rat cDNA clones coding for the autoantigen SS-B/La: detection of species-specific variations.";
 RT Gene 126:265-268 (1993).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' terminus of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including crRNA and 4.5S, 5S, 7S, and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: Phosphorylated (By similarity).
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; X67859; CAA48043.1; -
 DR PIR; JCI494; JCI494.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; xrm; 1.
 DR PRINTS; PR00302; LUPUSLA.

DT 01-JUL-1989 (Rel. 11, last sequence update)
 DT 15-MAR-2004 (Rel. 43, last annotation update)
 DE lupus la protein (Sjogren syndrome type B antigen) (SS-B) (La
 DE ribonucleoprotein) (La autoantigen).
 GN SSB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89202037; PubMed=2468131;
 RA Chan E.K.L., Sullivan K.F., Tan E.M.;
 RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus
 RT sequences for RNA-binding.";
 RL Nucleic Acids Res. 17:2235-2244(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89053970; PubMed=1192525;
 RA Chambers J.C., Kennan D., Martin B.J., Keene J.D.;
 RT "Genomic structure and amino acid sequence domains of the human La
 RT autoantigen.";
 RL J. Biol. Chem. 263:18043-18051(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta; and Skeletal muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uddin T.B., Tomihata S., Abramson R.D., Mullen S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Heltón E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallie D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE OF 54-408 FROM N.A.
 RX MEDLINE=88199081; PubMed=2452201;
 RA Sturges A.D., Peterson M.G., McNeillage L.J., Whittingham S.,
 RA Coppel R.S.;
 RT "Characteristics and epitope mapping of a cloned human autoantigen
 RT La.";
 RL J. Immunol. 140:3212-3218(1988).
 RN [5]
 RP SEQUENCE OF 54-97 FROM N.A.
 RX MEDLINE=85166283; PubMed=3856688;
 RA Chambers J.C., Keene J.D.;
 RT "Isolation and analysis of cDNA clones expressing human lupus La
 RT antigen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2115-2119(1985).
 RN [6]
 RP FUNCTION.
 RX MEDLINE=89251617; PubMed=2470590;
 RA Gottlieb E., Steltz J.A.;
 RT "Function of the mammalian La protein: evidence for its action in
 RT transcription termination by RNA polymerase III.";
 RL EMO J. 8:851-861(1989).
 RN [7]
 RP PHOSPHORYLATION.
 RX MEDLINE=97207017; PubMed=9054510;
 RA Fan H., Sakulich A.L., Goodier J.L., Zhang X., Qin J., Marate R.J.;

RT "Phosphorylation of the human La antigen on serine 366 can regulate
 RT recycling of RNA polymerase III transcription complexes.";
 RL Cell 88:707-715(1997).
 RN [8]
 RP INTERACTION WITH DDX15.
 RX MEDLINE=22346009; PubMed=12458796;
 RA Fournoux M.A., Kolkman M.J.M., Van der Heijden A., De Jong A.S.,
 RA Van Veenrooy W.J., Pruijn G.J.M.;
 RT "The human La (SS-B) autoantigen interacts with DDX15/hPrp43, a
 RT putative DEAD-box RNA helicase.";
 RL RNA 8:1428-1443(2002).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' termin of virtually all nascent
 CC polymerase III transcripts. It is associated with precursor forms
 CC of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S,
 CC and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE
 CC C-TERMINAL PART OF THE PROTEIN.
 CC -1- PTM: The N-terminus is blocked.
 CC -1- DISEASE: Sera from patients with systemic lupus erythematosus
 CC often contain antibodies that react with the normal cellular
 CC La protein as if this antigen was foreign.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; X13697; CA31985.1; -;
 DR EMBL; J04205; AAA51885.1; -;
 DR EMBL; BC001289; AA01289.1; -;
 DR EMBL; BC020818; AA020818.1; -;
 DR PIR; A31888; A31888.
 DR Genew; H8NC; 11316; SSB.
 DR MIM; 109090; -;
 DR GO; GO:0030529; C:ribonucleoprotein complex; TAS.
 DR GO; GO:0003729; F:mRNA binding; TAS.
 DR GO; GO:0000049; F:RNA binding; TAS.
 DR GO; GO:0008334; P:histone mRNA metabolism; TAS.
 DR GO; GO:0006400; P:RNA modification; TAS.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dom.
 DR InterPro; IPR00504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 DR Systemic lupus erythematosus; RNA-binding; Phosphorylation;
 KW Nuclear protein.
 FT DOMAIN 111 187 RNA-BINDING (RRM).
 FT MOD RES 366 366 PHOSPHORYLATION (BY CK2).
 SO SEQUENCE 408 AA; 46837 MW; EC153C15F9187FC4 CRC64;
 Query Match 67.9%; Score 57; DB 1; Length 408;
 Best Local Similarity 66.7%; Pred. No. 0.02;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 2 ERATIRVEYFGDF 16
 DB 14 EAKICQIEYFGDF 28
 RESULT 4

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:01:41 ; Search time 6.43575 Seconds
(without alignments)
129,452 Million cell updates/sec

Title: US-09-836-073-19

Perfect score: 84
Sequence: 1 QERAIRQVEYFGDF 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	92.9	390	1 LA_DROME	P40796 drosophila
2	57	67.9	404	1 LA_BOVIN	P10881 bos taurus
3	57	67.9	408	1 LA_HUMAN	P05455 homo sapien
4	57	67.9	415	1 LA_MOUSE	P32067 mus musculu
5	57	67.9	415	1 LA_RAT	P38656 rattus norv
6	56	66.7	383	1 LA_AEDEL	Q26457 aedes albop
7	48	57.1	427	1 LAB_XENLA	P28049 xenopus lae
8	48	57.1	428	1 LAB_XENLA	P28048 xenopus lae
9	45	54.8	298	1 LAH1_SCHPO	P87058 schizosacch
10	44	52.4	315	1 HO2_FAT	P33711 rattus norv
11	43	51.2	466	1 SNO9_YEAST	P25567 saccharomyc
12	43	51.2	569	1 VP45_ARATH	Q49048 arabidopsis
13	42	50.0	275	1 LAH1_YEAST	P33399 saccharomyc
14	42	50.0	398	1 YXGN_ECOS7	P58219 escherichia
15	42	50.0	398	1 YXGN_ECOLI	P39338 escherichia
16	41	48.8	500	1 SYH_EHILLO	Q98760 rhizobium l
17	40.5	48.2	506	1 TWMO_PETOR	Q43033 petroselinu
18	40	47.6	304	1 Y687_HAERI	P71356 haemophilus
19	40	47.6	373	1 TYRA_ECOLI	P07023 escherichia
20	40	47.6	634	1 IDUA_MOUSE	P18441 mus musculu
21	39	46.4	205	1 VADI_TREPA	O84411 treponema p
22	39	46.4	253	1 SOU_TREPA	O83326 treponema p
23	39	46.4	317	1 PSTB_XYLUFA	Q9PBD0 xyloella fas
24	39	46.4	312	1 HO2_RABIT	P43242 oryctolagus
25	39	46.4	315	1 HO2_MOUSE	O70252 mus musculu
26	39	46.4	315	1 HO2_HUMAN	O96651 homo sapien
27	39	46.4	368	1 PERQ_LACDE	O96651 lactobacilli
28	39	46.4	368	1 PERQ_LACDE	P46545 lactobacilli
29	39	46.4	532	1 IPA7_SHIFL	P18014 shigella fl
30	39	46.4	602	1 EXSA_BUCAL	P57530 buchnera ap
31	38	45.2	625	1 BGAL_LACSK	O48846 lactobacilli
32	37	44.0	216	1 POLG_PMDVC	P03309 foot-and-mo
33	37	44.0	230	1 POLG_PMDVC	P03307 foot-and-mo

34	37	44.0	234	1 POLG_PMDVI	P03310 foot-and-mo
35	37	44.0	281	1 Y818_PYPAB	Q8291 pyrobaculum
36	37	44.0	363	1 YCDM_ECOS7	Q8291 escherichia
37	37	44.0	363	1 YCDM_ECOLI	P75898 escherichia
38	37	44.0	550	1 SYR_MYCLE	P45840 mycobacteri
39	37	44.0	570	1 VP45_HUMAN	O92977 homo sapien
40	37	44.0	570	1 VP45_MOUSE	P97390 mus musculu
41	37	44.0	570	1 VP45_RAT	O08700 rattus norv
42	37	44.0	611	1 VAP2_PLAFA	Q03498 plasmodium
43	37	44.0	781	1 APE2_SULTO	O97496 sulfolobus
44	37	44.0	900	1 GDNH_CLOTM	P16218 clostridium
45	37	44.0	905	1 CHD2_SCHPO	O74787 schizosacch

ALIGNMENTS

RESULT 1
LA_DROME STANDARD; PRT; 390 AA.
ID LA_DROME
AC P40796; Q24375; Q9V1N2;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
GN LA OR CG10922.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S; TISSUE=Ovary;
RX MEDLINE=94309632; PubMed=8035794;
RA Bai C., Li Z., Tollas P.P.;
RT "Developmental characterization of a Drosophila RNA-binding protein homologous to the human systemic lupus erythematosus-associated La/SS-B autoantigen."
RT Mol. Cell. Biol. 14:5123-5129 (1994).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Adamas W.D., Baker J., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-U., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Baer A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintinas S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Deitcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitani M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

This Page Blank (uspto)

T00677
 hypothetical protein At2g43970 [imported] - Arabidopsis thaliana
 N/Alternate names: hypothetical protein F6E13.10
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 23-Mar-2001
 C/Accession: T00677, G84872
 R/Ronday, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, June 1998
 A/Description: Arabidopsis thaliana chromosome II BAC F6E13 genomic sequence.
 A/Reference number: Z14180
 A/Accession: T00677
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-529 <R0U>
 A/Cross-references: EMBL:AC004005; NID:G3212846; PID:G3212854
 A/Experimental source: Cultivar Columbia
 R/Lin, X.; Kaul, S.; Ronday, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; Neus, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A/Reference number: A84420; MUID:20083487; PMID:10617197
 A/Accession: G84872
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-529 <STO>
 A/Cross-references: GB:AE002093; NID:G3212854; PIDN:AAC23405.1; GSPDB:GN00139
 A/Map position: 2
 A/Introns: 200/3; 228/1; 259/3; 283/3; 310/2; 344/3; 367/3; 421/3

Query Match 56.0%; Score 47; DB 2; Length 529;
 Best Local Similarity 72.7%; Pred. No. 3.1;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IIR0VEYFGD 15
 | : | | | | | |
 Db 197 IVN0VEYFSD 207

RESULT 11
 T38937
 rna binding protein - fission yeast (Schizosaccharomyces pombe)
 C/Species: Schizosaccharomyces pombe
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C/Accession: T38937
 R/Baddock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, April 1997
 A/Reference number: Z21818
 A/Accession: T38937
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-298 <BAD>
 A/Cross-references: EMBL:Z24864; PIDN:CA08173.1; GSPDB:GN00066; SPDB:SPAC57A10.100
 A/Experimental source: strain 972h-; cosmid c57A10
 C/Genetics:
 A/Gene: SPDB:SPAC57A10.100
 A/Map position: 1
 A/Introns: 72/1

Query Match 54.8%; Score 46; DB 2; Length 298;
 Best Local Similarity 50.0%; Pred. No. 2.5;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 ERAIIR0VEYFGD 15
 | : | | | | | |
 Db 64 EAEVLK0VEYFSD 77

RESULT 12
 T43542
 RNA-binding protein Lal homolog - fission yeast (Schizosaccharomyces pombe)

N/Alternate names: La autocatalytic; ribonucleoprotein La homolog
 C/Species: Schizosaccharomyces pombe
 C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000
 C/Accession: T43542; T43325
 R/Van Horn, D.J.; Yoo, C.J.; Xue, D.; Shi, H.; Molin, S.L. RNA 3, 1434-1443, 1997
 A/Title: The La protein in Schizosaccharomyces pombe: a conserved yet dispensable phospho
 A/Reference number: Z22560; MUID:98067398; PMID:9404894
 A/Accession: T43542
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-298 <VAN>
 A/Cross-references: EMBL:AF022949; PIDN:AAB82145.1
 R/Utsami, R. submitted to the EMBL Data Library, February 1998
 A/Description: Screening of S. pombe cDNA library using E. coli defective in signal tran
 A/Reference number: Z22428
 A/Accession: T43325
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-298 <UTS>
 A/Cross-references: EMBL:AB011371; PIDN:BAA24981.1
 C/Genetics:
 A/Gene: slal
 A/Function:
 A/Description: the binding of the La protein to rRNA precursors is required for the endo
 C/Keywords: phosphoprotein; RNA binding

Query Match 54.8%; Score 46; DB 2; Length 298;
 Best Local Similarity 50.0%; Pred. No. 2.5;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 ERAIIR0VEYFGD 15
 | : | | | | | |
 Db 64 EAEVLK0VEYFSD 77

RESULT 13
 H83743
 ABC transporter (permease) BH0752 [imported] - Bacillus halodurans (strain C-125)
 C/Species: Bacillus halodurans
 C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C/Accession: H83743
 R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira Nucleic Acids Res. 28, 4317-4331, 2000
 A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A/Reference number: A83650; MUID:20512582; PMID:11058132
 A/Accession: H83743
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-642 <STO>
 A/Cross-references: GB:AP001509; GB:BA000004; NID:G10173176; PIDN:BA04471.1; GSPDB:GN000
 A/Experimental source: strain C-125
 C/Genetics:
 A/Gene: BH0752

Query Match 54.2%; Score 45.5; DB 2; Length 642;
 Best Local Similarity 60.0%; Pred. No. 7;
 Matches 9; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 1 ORAIIR0VEYFGD 15
 | : | | | | | |
 Db 471 ORKALVLR-EYFSD 484

RESULT 14
 A35199
 heme oxygenase (decyclizing) (EC 1.14.99.3) 2 [similarity] - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C/Accession: A35199; A29922
 R/Rotenberg, M.O.; Maines, M.D. J. Biol. Chem. 265, 7501-7506, 1990

A:Cross-references: GB:X67859; NID:g55778; PIDN:CAA48043.1; PID:g55779
 A:Experimental source: liver
 A:Comment: This protein associates with a variety of small RNA molecules, most of which
 act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:112-178/Domain: ribonucleoprotein repeat homology <RRM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:127-415/Domain: phosphorylated #status predicted <PHY>

Query Match 67.9%; Score 57; DB 1; Length 415;
 Best Local Similarity 66.7%; Pred. No. 0.042;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 2 ERAIIRQVEYFGDF 16
 Db 14 EAKICHQIEYFGDF 28

RESULT 6
 T30953
 hypothetical protein C44E4.4 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000
 C:Accession: T30953
 R:Sammons, L.; Wohlmann, P.; Gillam, B.
 submitted to the EMBL Data Library, August 1999
 A:Description: The sequence of *C. elegans* cosmid C44E4.
 A:Reference number: Z20945
 A:Accession: T30953
 A:Status: preliminary; translated from GB/EMBL/DDBF
 A:Molecule type: DNA
 A:Residues: 1-396 <SAM>
 A:Cross-references: EMBL:AF003140; PIDN:AAB54169.1
 A:Experimental source: strain Bristol N2; clone C44E4
 C:Genetics:
 A:Map position: 1
 A:Insertions: 45/1; 114/3
 A:Note: C44E4.4
 C:Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology

Query Match 59.5%; Score 50; DB 2; Length 396;
 Best Local Similarity 57.1%; Pred. No. 0.67;
 Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 2 ERAIIRQVEYFGD 15
 Db 13 DQKITKQLEYFPGN 26

RESULT 7
 S33817
 ribonucleoprotein la.B - African clawed frog
 N:Alternate names: autoantigen SS-B/la; ribonucleoprotein SS-B
 C:Species: *Xenopus laevis* (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33817; S28544
 R:Scherly, D.; Stutz, F.; Lin-Marq, N.; Clarkson, S.G.
 J. Mol. Biol. 231, 196-204, 1993
 A:Title: La proteins from *Xenopus laevis*. cDNA cloning and developmental expression.
 A:Reference number: S33817; MUID:93287095; PMID:8510143
 A:Accession: S33817
 A:Molecule type: mRNA
 A:Residues: 1-427 <SCH>
 A:Cross-references: EMBL:X68818; NID:g64875; PIDN:CAA48716.1; PID:g64876
 C:Comment: This protein associates with a variety of small RNA molecules, most of which
 act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:111-177/Domain: ribonucleoprotein repeat homology <RRM>
 F:112-117/Region: RNA-binding RNP2 motif
 F:150-157/Region: RNA-binding RNP1 motif

F:227-427/Domain: phosphorylated #status predicted <PHY>

Query Match 57.1%; Score 48; DB 1; Length 427;
 Best Local Similarity 72.7%; Pred. No. 1.6;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 5 IIRQVEYFGD 15
 Db 16 ICEQIEYFGD 26

RESULT 8
 S33818
 ribonucleoprotein la.A - African clawed frog
 N:Alternate names: autoantigen SS-B/la; ribonucleoprotein SS-B
 C:Species: *Xenopus laevis* (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33818; S28545
 R:Scherly, D.; Stutz, F.; Lin-Marq, N.; Clarkson, S.G.
 J. Mol. Biol. 231, 196-204, 1993
 A:Title: La proteins from *Xenopus laevis*. cDNA cloning and developmental expression.
 A:Reference number: S33817; MUID:93287095; PMID:8510143
 A:Accession: S33818
 A:Molecule type: mRNA
 A:Residues: 1-428 <SCH>
 A:Cross-references: EMBL:X68817; NID:g64873; PIDN:CAA48715.1; PID:g64874
 A:Comment: This protein associates with a variety of small RNA molecules, most of which
 act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:112-178/Domain: ribonucleoprotein repeat homology <RRM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:228-428/Domain: phosphorylated #status predicted <PHY>

Query Match 57.1%; Score 48; DB 1; Length 428;
 Best Local Similarity 72.7%; Pred. No. 1.6;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IIRQVEYFGD 15
 Db 17 ICEQIEYFGD 27

RESULT 9
 T04453
 hypothetical protein PAD11.80 - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999
 C:Accession: T04453
 R:Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansgorge, W.; Hohenseel, J.; Mewes, H.
 submitted to the Protein Sequence Database, April 1998
 A:Reference number: Z15360
 A:Accession: T04453
 A:Molecule type: DNA
 A:Residues: 1-483 <BEV>
 A:Cross-references: EMBL:AL022537
 A:Experimental source: cultivar Columbia; BAC clone PAD11
 C:Genetics:
 A:Map position: 4
 A:Insertions: 17/3; 44/1; 94/1; 141/3; 206/2; 247/2; 345/3; 395/1; 434/2
 A:Note: PAD11.80

Query Match 57.1%; Score 48; DB 2; Length 483;
 Best Local Similarity 53.8%; Pred. No. 1.9;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 RAIIRQVEYFGD 15
 Db 12 KTVLRQVEYFGD 24

RESULT 10

Best Local Similarity 100.0%; Pred. No. 8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QERAIROVEYFGDF 15
|||||
Db 50 QERAIROVEYFGDF 64

RESULT 3

S03849

ribonucleoprotein Ia - bovine

N/Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B

C/Species: Bos primigenius taurus (cattle)

C/Date: 31-Dec-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C/Accession: S03849

R/Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.

Nucleic Acids Res. 17, 2233-2244, 1989

A/Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences

A/Reference number: S03848; PMID:89202037; PMID:2468131

A/Accession: S03849

A/Molecule type: mRNA

A/Residues: 1-404 <CHA>

A/Cross-references: EMBL:X13698; NID:g755; PIDN:CAA31986.1; PID:g756

A/Note: Part of this sequence was confirmed by protein sequencing

C/Comment: This protein associates with a variety of small RNA molecules, most of which

act as a transcription termination factor.

C/Superfamily: ribonucleoprotein Ia; ribonucleoprotein repeat homology

C/Keywords: blocked amino end; phosphoprotein; RNA binding

F/112-178/Domain: ribonucleoprotein repeat homology <RMM>

F/113-118/Region: RNA-binding RNP2 motif

F/151-158/Region: RNA-binding RNP1 motif

F/228-404/Domain: phosphorylated #status predicted <PHY>

Query Match 67.9%; Score 57; DB 1; Length 404;

Best Local Similarity 66.7%; Pred. No. 0.041;

Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ERAIROVEYFGDF 16
|:|||||
Db 14 EAKICHOIEYFGDF 28

RESULT 4

A31888

ribonucleoprotein Ia - human

N/Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B; Sjogren syndrome antigen

C/Species: Homo sapiens (man)

C/Date: 21-May-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C/Accession: A31888; S03848; A22956; A61051; S11013; I55553; I70205; I70206; A31273

R/Chan, E.K.L.; Kenan, D.; Martin, B.J.; Keene, J.D.

J. Biol. Chem. 263, 18043-18051, 1988

A/Title: Genomic structure and amino acid sequence domains of the human Ia autoantigen.

A/Reference number: A31888; PMID:89053970; PMID:3192555

A/Accession: A31888

A/Molecule type: mRNA

A/Residues: 1-408 <CHA>

A/Cross-references: GB:J04205; NID:g178686; PIDN:AAA51885.1; PID:g178687

R/Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.

Nucleic Acids Res. 17, 2233-2244, 1989

A/Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences

A/Reference number: S03848; PMID:89202037; PMID:2468131

A/Accession: S03848

A/Molecule type: mRNA

A/Residues: 1-408 <CH2>

A/Cross-references: EMBL:X13697; NID:g36414; PIDN:CAA31985.1; PID:g36415

R/Chan, E.K.L.; Keene, J.D.

Proc. Natl. Acad. Sci. U.S.A. 82, 2115-2119, 1985

A/Title: Isolation and analysis of cDNA clones expressing human lupus Ia antigen.

A/Reference number: A22956; PMID:85166283; PMID:3856888

A/Accession: A22956

A/Molecule type: mRNA

A/Residues: 45-97, 'LK' <CH3>

A/Cross-references: GB:J04205

A/Note: this sequence has been revised in reference A31888

R/Nyman, U.; Ringertz, N.R.; Petersson, I.

Immunol. Lett. 22, 65-72, 1989

A/Title: Demonstration of an amino terminal Ia epitope recognized by human anti-Ia sera.

A/Reference number: A61051; PMID:89379261; PMID:2476379

A/Accession: A61051

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-19, 'E', '21-47' <YMM>

R/Sturges, A.D.; Peterson, M.G.; McNeillage, L.J.; Whittingham, S.; Coppel, R.L.

J. Immunol. 140, 3212-3218, 1988

A/Title: Characteristics and epitope mapping of a cloned human autoantigen Ia.

A/Reference number: S11013; PMID:88199081; PMID:2452201

A/Accession: S11013

A/Molecule type: mRNA

A/Residues: 'E', '55-287', 'V', '289-408' <STU>

A/Cross-references: EMBL:M20328; NID:g337456; PIDN:AAA6557.1; PID:g337457

R/Kohaka, H.; Yamamoto, K.; Fujii, H.; Miura, H.; Miyasaka, N.; Nishio, K.; Miyamoto, J.

Clin. Invest. 85, 1566-1574, 1990

A/Title: Fine epitope mapping the human SS-B/La protein: Identification of a distinct at

A/Reference number: I55553; PMID:90237237; PMID:1692037

A/Accession: I55553

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Cross-references: GB:M35261; NID:g338492; PIDN:AAA36652.1; PID:g338495

A/Accession: I70205

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 174-224 <RR2>

A/Cross-references: GB:M35263; NID:g338492; PIDN:AAA36653.1; PID:g338496

A/Accession: I70206

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 279-342 <RR3>

A/Cross-references: GB:M35262; NID:g338493; PIDN:AAA36654.1; PID:g338497

C/Comment: This protein associates with a variety of small RNA molecules, most of which

act as a transcription termination factor.

C/Genetics:

A/Status: GDB:SSB

A/Cross-references: GDB:125359; OMIM:109090

A/Map position: 2

A/Introns: 22/3; 57/2; 115/3; 151/3; 185/2; 209/2; 223/3; 264/3; 380/2

C/Superfamily: ribonucleoprotein Ia; ribonucleoprotein repeat homology

C/Keywords: phosphoprotein; RNA binding

F/112-178/Domain: ribonucleoprotein repeat homology <RMM>

F/113-118/Region: RNA-binding RNP2 motif

F/151-158/Region: RNA-binding RNP1 motif

F/228-408/Domain: phosphorylated #status experimental <PHY>

Query Match 67.9%; Score 57; DB 1; Length 408;

Best Local Similarity 66.7%; Pred. No. 0.041;

Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ERAIROVEYFGDF 16
|:|||||
Db 14 EAKICHOIEYFGDF 28

RESULT 5

JC1494

ribonucleoprotein Ia - rat

N/Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B

C/Species: Rattus norvegicus (Norway rat)

C/Date: 24-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C/Accession: JC1494; S25145

R/Semel, I.; Troester, H.; Bartsch, H.; Schwemmler, M.; Igloi, G.L.; Bachmann, M.

Gene 126, 265-268, 1993

A/Title: Isolation of rat cDNA clones coding for the autoantigen SS-B/La: Detection of

A/Reference number: JC1494; PMID:93246255; PMID:7916708

A/Accession: JC1494

A/Molecule type: mRNA

A/Residues: 1-415 <SEM>

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:33:11 / Search time 12.514 Seconds
(without alignments)
122.988 Million cell updates/sec

Title: US-09-836-073-19

Perfect score: 84
Sequence: 1 QERAIIRQVEYFQDF 16

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

PIR.78:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	92.9	390	2	la/SS-B homolog D-
2	78	92.9	390	2	ribonucleoprotein
3	57	67.9	404	1	ribonucleoprotein
4	57	67.9	408	1	ribonucleoprotein
5	57	67.9	415	1	ribonucleoprotein
6	50	59.5	396	2	ribonucleoprotein
7	48	57.1	427	1	ribonucleoprotein
8	48	57.1	428	1	ribonucleoprotein
9	48	57.1	483	2	hypothetical prote
10	47	56.0	529	2	hypothetical prote
11	46	54.8	298	2	RNA binding protei
12	46	54.8	298	2	ABC transporter (p
13	45.5	54.2	642	2	heme oxygenase (de
14	44	52.4	315	1	hypothetical prote
15	43	51.2	466	2	hypothetical prote
16	43	51.2	469	2	hypothetical prote
17	43	51.2	569	2	vacuolar protein s
18	43	51.2	569	2	vacuolar protein s
19	43	51.2	846	2	general secretion
20	42	50.0	275	2	RNA-binding protei
21	42	50.0	393	2	hypothetical 44.4
22	42	50.0	398	2	hypothetical prote
23	42	50.0	407	2	hypothetical prote
24	42	50.0	788	2	hypothetical prote
25	41.5	49.4	334	2	probable proteinas
26	41	48.8	421	2	hypothetical prote
27	41	48.8	483	2	hypothetical prote
28	41	48.8	520	2	hypothetical prote
29	40.5	48.2	506	2	trans-cinnamate 4-

30	40.5	48.2	3078	2	T28432	variant-specific s
31	40	47.6	230	2	C84410	hypothetical prote
32	40	47.6	304	2	H64156	hypothetical prote
33	40	47.6	373	1	KMECD	chorismate mutase
34	40	47.6	373	2	A10832	prephenate dehydro
35	40	47.6	373	2	C85906	chorismate mutase-
36	40	47.6	373	2	G91061	chorismate mutase-
37	40	47.6	384	2	H96829	probable RNA-bindi
38	40	47.6	391	2	H89777	capsular polysacch
39	40	47.6	634	1	A55683	L-iduronidase (EC
40	40	47.6	662	2	S55387	hemocyanin precurs
41	40	47.6	736	1	C69307	conserved hypochet
42	39.5	47.0	335	2	A75179	probable proteinas
43	39	46.4	206	2	A71326	probable V-type At
44	39	46.4	253	2	B71346	probable Spo0V reg
45	39	46.4	267	2	D82593	phosphate ABC tran

ALIGNMENTS

RESULT 1

A53773
la/SS-B homolog D-1a - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
C/Accession: A53773
R/Bat, C./ Li, Z.; Tollas, P.P.
Mol. Cell. Biol. 14, 5123-5129, 1994
A/Title: Developmental characterization of a Drosophila RNA-binding protein homologous
A/Reference number: A53773; MUID:94309632; PMID:8035794
A/Accession: A53773
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-390 <BAT>
A/Cross-references: GB:007652, NID:9464019, PIDN:AAA20518.1, PID:9464020
C/Genetics:
A/Gene: FlyBase:la
A/Cross-references: FlyBase:FBgn0011638
C/Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology
C/Keywords: leucine zipper, RNA binding

Query Match 92.9%; Score 78; DB 2; Length 390;
Best Local Similarity 100.0%; Pred. No. 8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QERAIIRQVEYFQDF 15
Db 50 QERAIIRQVEYFQDF 64

RESULT 2

A53781
ribonucleoprotein la - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 20-Sep-1999
C/Accession: A53781
R/Yoo, C.J.; Molin, S.L.
Mol. Cell. Biol. 14, 5412-5424, 1994
A/Title: la proteins from Drosophila melanogaster and Saccharomyces cerevisiae: a yeast
A/Reference number: A53781; MUID:94309661; PMID:8035818
A/Accession: A53781
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-390 <YOO>
A/Cross-references: GB:L32988, NID:9488469, PID:9488470
C/Genetics:
A/Gene: FlyBase:la
A/Cross-references: FlyBase:FBgn0011638
C/Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology
C/Keywords: RNA binding

Query Match 92.9%; Score 78; DB 2; Length 390;

08YOB2

ID 08YOB2 PRELIMINARY; PRT; 451 AA.

AC 08YOB2; PRELIMINARY; PRT; 451 AA.

DT 01-MAR-2002 (TEMBLrel. 20, Created)

DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)

DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)

DE UDP-N-acetylglucosamine pyrophosphorylase.

GN ALR3921.

OS Anabaena sp. (strain PCC 7120).

OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.

OX NCBI_TaxID=103690;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=2159285; PubMed=11759840;

RA Kaneo T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,

RA Kishida Y., Kohara M., Matsuno M., Matsuno A., Muraki A.,

RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,

RA Yasuda M., Tabata S.,

RT "Complete genomic sequence of the filamentous nitrogen-fixing

cyanobacterium Anabaena sp. strain PCC 7120."

RL DNA Ref. 8:205-213(2001).

EMBL; AP003594; BAB75620.1; -.

DR GO; GO:0003977; P:UDP-N-acetylglucosamine diphosphorylase act. .; IEA.

DR GO; GO:0009058; P:biogenesis; IEA.

DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.

DR InterPro; IPR001451; Hexapep_transf.

DR InterPro; IPR005835; NTP_transferase.

DR Pfam; PF00132; hexapep; 6.

DR Pfam; PF00483; NTP_transferase; 1.

DR TIGRPFAM; TIGR01173; glmU; 1.

KW COMPLETE proteome.

SQ SEQUENCE 451 AA; 49191 MW; B1DEB3A6F2170FA7 CRC64;

Query Match 46.7%; Score 42; DB 16; Length 451;

Best Local Similarity 53.3%; Pred. No. 65;

Matches 8; Conservative. 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 LEASTIROEYFGD 18

DB 187 LEANNNAQKEYVLTDA 201

RESULT 15

ID 004697

AC 004697

DT 01-JUL-1997 (TEMBLrel. 04, Created)

DT 01-JUL-1997 (TEMBLrel. 04, Last sequence update)

DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)

DE DNA-binding protein PD2.

OS Pisum sativum (Garden pea).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosid I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.

OX NCBI_TaxID=3888;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Alaska; TISSUE=leaf;

RA Sato N., Kazuno A.A., Ohta N., Ohshima K.,

RT "Identification of a novel family of DNA-binding proteins with two AT-

hook motifs from pea."

RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

EMBL; X98740; CAA67292.1; -.

DR PIR; T06586; T06586.

DR GO; GO:0003677; P:DNA binding; IEA.

KW DNA-binding.

SQ SEQUENCE 632 AA; 69498 MW; 9F744E227CD08717 CRC64;

Query Match 46.7%; Score 42; DB 10; Length 632;

Best Local Similarity 50.0%; Pred. No. 95;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 LEASTIROEYFGD 17

DB 161 VBATVSDRYTGD 174

Search completed: September 10, 2004, 18:00:17

Job time : 44.229 secs

DR ProDom; PD000001; Prot Kinase; 1.
 DR SMART; SM00192; LDR; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR SMART; SM00219; TYKC; 1.
 DR PROSITE; PS01209; LDR_1; 1.
 DR PROSITE; PS0068; LDR_2; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
 DR ATP-Binding; Transferase.
 FT NON TER 1
 SQ SEQUENCE 1161 AA; 124852 MW; 601210E788E8AA41 CRC64;

Query Match 47.2%; Score 42.5; DB 13; Length 1161;
 Best Local Similarity 43.5%; Pred. No. 1.5e+02;
 Matches 10; Conservative 4; Mismatches 4; Indels 5; Gaps 1;

QY 1 VSKLEASTROEYFGD 18
 DB 530 LSKIRSTIMTDNPNNGYFGKA 552

RESULT 11

088YN2 PRELIMINARY; PRT; - 160 AA.
 AC 088YN2;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-OCT-2002 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Ribosomal-protein-alanine N-acetyltransferase (EC 2.3.1.128).
 CN RM11 OR LP 0718.
 OS Lactobacillus plantarum.
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 OC Lactobacillus.
 OX NCBI_TaxID=1590;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCIMB 8826 / WCFS1.
 RX MEDLINE=22480296; PubMed=12566566;
 RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
 RA Kijpers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
 RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
 RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
 RA De Vos W.M., Siezen R.J.;
 RT "Complete genome sequence of Lactobacillus plantarum WCFS1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
 DR EMBL; AL935253; CAD63319.1; -
 DR GO; GO:0008415; F:acetyltransferase activity; IEA.
 DR GO; GO:0008080; F:N-acetyltransferase activity; IEA.
 DR GO; GO:0008999; F:ribosomal-protein-alanine N-acetyltransferase. . .; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR InterPro; IPR001182; GCSAcetyl_trans.
 DR Pfam; PF00583; Acetyltransf. 1.
 KW Acyltransferase; Transferase; Complete proteome.
 SQ SEQUENCE 160 AA; 18734 MW; 3FD115B095A526 CRC64;

Query Match 46.7%; Score 42; DB 16; Length 160;
 Best Local Similarity 53.8%; Pred. No. 20;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 5 EASTIROEYFGD 17
 DB 131 EKTGIRKRYPGD 143

RESULT 12

088XT9 PRELIMINARY; PRT; 308 AA.
 AC 088XT9;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE Glycosyltransferase.
 GN MM1127.
 OS Methanosarcina mazel (Methanosarcina frisia).
 OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
 OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2209;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Geel / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;
 RX MEDLINE=22108827; PubMed=12125824;
 RA Depgenmeier U., Johann A., Hartisch T., Merkl R., Schmitz R.A.,
 RA Martinek-Artas R., Henne A., Wierzer A., Benner S., Jacobi C.,
 RA Brueggemann H., Lenard T., Christmann A., Boemecke M., Steckel S.,
 RA Bhattacharyya A., Lykdis A., Overbeek R., Klenk H.-P., Gunsalus K.P.,
 RA Fritz H.-J., Gottschalk G.;
 RT "The genome of Methanosarcina mazel: evidence for lateral gene
 transfer between Bacteria and Archaea.";
 RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
 DR EMBL; AE013341; AAM30823.1; -
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR InterPro; IPR01173; Glyco trans 2.
 DR Pfam; PF00535; Glycos transf 2; 1.
 KW Transferase; Complete proteome.
 SQ SEQUENCE 308 AA; 35128 MW; 0C01D9C354DA40DE CRC64;

Query Match 46.7%; Score 42; DB 17; Length 308;
 Best Local Similarity 50.0%; Pred. No. 42;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 VSKLEASTROEYFG 16
 DB 105 IGLDADTLEENYFG 120

RESULT 13

086119 PRELIMINARY; PRT; 358 AA.
 AC 086119;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Hypothetical protein.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RX MEDLINE=22092622; PubMed=12097910;
 RA Gloeckner G., Richinger L., Szafranski K., Pachebat J., Dear P.,
 RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
 RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
 RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
 RL Nature 418:79-85(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA Baumgart C.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC117176; AAO52103.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 358 AA; 39920 MW; 021273F80378AF7C CRC64;

Query Match 46.7%; Score 42; DB 5; Length 358;
 Best Local Similarity 53.3%; Pred. No. 50;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 3 KLEASTROEYFGD 17
 DB 64 KLISKTIKSKTFGS 78

RESULT 14

```
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0006535; P:cysteine biosynthesis from serine; IEA.
DR GO; GO:0013343; P:cysteine biosynthesis via cystathione; IEA.
DR InterPro; IPR001926; B6_enzyme_beta.
DR InterPro; IPR000644; CBS_domain.
DR InterPro; IPR005857; Cysa_beta_synth.
DR InterPro; IPR001216; Cys_synthase_BS.
DR Pfam; PF00571; CBS; 2.
DR Pfam; PF00291; PALP; 1.
DR SMART; SM00116; CBS; 2.
DR TIGRFAMs; TIGR01137; cysa_beta; 1.
DR PROSITE; PS00301; CYS_SYNTHASE; 1.
KW Lyase.
SQ SEQUENCE 507 AA; 56025 MW; 1589165DEB7C8A54 CRC64;

Query Match
Best Local Similarity 47.8%; Score 43; DB 3; Length 507;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 SKLEASTIRQEVYFGDA 18
Db 351 SKLEASTIRQADVFGNA 367

RESULT 8
Q8E9H6 PRELIMINARY; PRT; 206 AA.
AC Q8E9H6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Conserved hypothetical protein.
GN SO4307.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_Taxid=70663;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NR-1;
RA MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer R.T., Tsapin R.J., Beaman M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Uppraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Uterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Nelson K.H., Fraser C.M.,
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RT Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AEO15863; AAN57276.1; -
DR TIGR; SO4307; -
DR InterPro; IPR007435; DUF484.
DR Pfam; PF04340; DUF484; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 206 AA; 24078 MW; D8935EAD40692P10 CRC64;

Query Match
Best Local Similarity 47.2%; Score 42.5; DB 16; Length 206;
Matches 10; Conservative 3; Mismatches 3; Indels 9; Gaps 1;

Qy 1 VSKLEASTIRQEVYFG 16
Db 118 ISVLBADVEMKAIWAERLKEYYFG 142

RESULT 9
Q7ZT10 PRELIMINARY; PRT; 401 AA.
AC Q7ZT10;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
```

```
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to sjogren syndrome antigen B (Autoantigen Ia).
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_Taxid=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC045392; AAHA5392.1; -
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR00630; Lupus_La_dom.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 401 AA; 46138 MW; 2EF032FDD3916291 CRC64;

Query Match
Best Local Similarity 47.2%; Score 42.5; DB 13; Length 401;
Matches 10; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

Qy 1 VSKLEASTIRQEVYFGD 17
Db 7 MSPLEKVAEOIEYFGD 24

RESULT 10
Q8JFV0 PRELIMINARY; PRT; 1161 AA.
AC Q8JFV0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 25, Last annotation update)
DE SI:D2107016.1 (Novel protein similar to vertebrate anaplastic lymphoma
DE kinase (ALK) and leukocyte tyrosine kinase receptor precursor (LTK or
DE TYK1)) (Fragment).
GN SI:D2107016.1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_Taxid=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL596022; CAD43463.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0004714; F:transmembrane receptor protein tyrosine kin. . .; IEA.
DR GO; GO:0004468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002011; RecepttyrkinetI.
DR InterPro; IPR002290; Set_tnr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00057; ldl_recept_a; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
```

```

RC STRAIN=ATCC 12228;
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
RA Chen Z., Wen Y.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AE016750; AAC05481.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 116 AA; 14002 MW; C8419BFA5B8A7EA5 CRC64;

Query Match      48.3%; Score 43.5; DB 16; Length 116;
Best Local Similarity 57.9%; Pred. No. 7.5;
Matches 11; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

Oy      1 VSKLEAST-IROEYFGDA 18
       :|::|||:|||
Db      78 ISKLRSFDERQVIYFFDA 96

RESULT 5
O9CGB9          PRELIMINARY; PRT; 381 AA.
ID O9CGB9
AC O9CGB9
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Sjogren syndrome antigen B.
GN SSB.
NC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Komio H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
RA Saito T., Okasaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner W., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glass C., King B., Kochiya H.,
RA Kuehl P., Lewis S., Matsumoto Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stenubi F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Bash G.,
RA Blake J., Botelli D., Bouunga N., Caminci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guestinich S., Hill D., Hofmann M., Hume D.A., Kamlaya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Saesaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyokaka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wyznali-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlschki S.,
RA Hayashizaki Y.;
RL "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK017822; BAB30957.1; -.
DR MGI; MGI:96423; SSB.
DR GO; GO:0005634; C:nucleus; IDA.
DR InterPro; IPRO02344; Lupus La.
DR InterPro; IPRO06630; Lupus La.dom.
DR InterPro; IPRO00504; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR PRINTS; PRO0302; LUPUSLA.
DR SMART; SM00715; LA; 1.
SQ SEQUENCE 381 AA; 43891 MW; 2E2DEF1452CF0F0E9 CRC64;

Query Match      48.3%; Score 43.5; DB 11; Length 381;
Best Local Similarity 55.6%; Pred. No. 29;
Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

Oy      1 VSKLEASTIRO-EYYFGD 17
       ::|||:|||
Db      10 MTALAKICHQILEYFGD 27
```

RESULT 6						
ID	Q08BTU4	PRELIMINARY;	PRT;	415 AA.		
AC	Q08BTU4					
DT	01-MAR-2003 (TREMBLrel_23, Created)					
DR	01-MAR-2003 (TREMBLrel_23, last sequence update)					
DT	01-OCT-2003 (TREMBLrel_25, last annotation update)					
DE	Sjogren syndrome antigen B.					
OS	Ssb.					
OC	Mus musculus (Mouse).					
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.					
OX	NCBI_TaxId=10090;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=NOD; TISSUE=Thymus;					
FX	MEDLINE=22354683; PubMed=12466851;					
RA	The FANTOM Consortium,					
RT	"Analysis of the mouse transcriptome based on functional annotation of					
RL	60,770 full-length cDNAs."					
RL	Nature 420:563-573 (2002).					
DR	EMBL; AK088677; BAC#0498.1; -.					
DR	MGI; MGI:98423; Ssb.					
DR	GO: GO:0005634; C:nucleus; IDA.					
DR	InterPro; IPR002344; Lupus_La.					
DR	InterPro; IPR006830; Lupus_La.dom.					
DR	InterPro; IPR000504; RNA_rec_mot.					
DR	Pfam; PF05383; La; 1.					
DR	Pfam; PF00076; rrm; 1.					
DR	PRINTS; PR00302; LUPUSLA.					
DR	SMART; SM00715; LA; 1.					
DR	SMART; SM00360; RRM; 1.					
DR	PROSITE; PS50102; RRM; 1.					
DR	PROSITE; PS00030; RRM_NRP_1; 1.					
SO	SEQUENCE 415 AA; 47657 MW; A7545C7686AC8363 CRC64;					
Query Match	48.3%;	Score 43.5;	DB 11;	Length 415;		
Best local similarity	55.6%;	Pred. No. 32;				
Matches 10; Conservative	2;	Mismatches 5;	Indels 1;	Gaps 1;		
Cy	1 VSKLEASTIRQ-EYFPGD 17					
D6	::::					
	10 WTALAKICHOIEYFGD 27					
RESULT 7						
ID	Q04073	PRELIMINARY;	PRT;	507 AA.		
AC	Q04073;					
DT	01-MAY-1999 (TREMBLrel_10, Created)					
DT	01-MAY-1999 (TREMBLrel_10, Last sequence update)					
DT	01-OCT-2003 (TREMBLrel_25, last annotation update)					
DE	Cytathionine beta-synthase (EC 4.2.1.22).					
GN	NHS3.					
OS	Saccharomyces cerevisiae (Baker's yeast).					
CC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;					
CC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.					
OX	NCBI_TaxID=4932;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	'STRAIN=X2180-1A';					
RA	Tezuka H., Mori T., Okumura Y., Kitabatake K., Tsamura Y.;					
RT	"Cloning of a gene suppression hypodrogen sulfide production by					
RT	Saccharomyces cerevisiae and its expression in a brewing yeast.";					
RL	ASBC Journal 50:130-133(1992).					
DR	EMBL; D16496; BAA03947.1; -.					
DR	HSSP; P35520; JUBO.					
DR	GO: GO:0005737; C:cyttoplasm; IEA.					
DR	GO: GO:0004123; F:cystathionine beta-synthase activity; IEA.					
DR	GO: GO:0016829; F:lyase activity; IEA.					


```

RESULT 2
O9SV58 ID O9SV58 PRELIMINARY; PRT; 365 AA.
AC O9SV58:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
GN F25G13.200 OR AT4G13110.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Eumariophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
CX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W.,
RA Mayer K.F.X., Lemcke K., Schueller C.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Pohl T., Weizenegger T., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL079349; CAB45510.1; -.
DR EMBL; AL161535; CAB78353.1; -.
DR PIR; T10213; T10213.
DR InterPro; IPR005607; BSD.
DR Pfam; PF03909; BSD; 1.
DR PROSITE; PS50858; BSD; 1.
KW Hypothetical protein.
SQ SEQUENCE 365 AA; 41867 MW; 47BDCCA1B18F43D2 CRC64;

Query Match 48.9%; Score 44; DB 10; Length 365;
Best Local Similarity 38.9%; Pred. No. 22;
Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0

Oy 1 VSKLEASTIRQRYFGDA 18
   : : : : : : : :
Db 8 ILRCSCTRREYFFHDA 25

RESULT 3
O9VZP8 ID O9VZP8 PRELIMINARY; PRT; 512 AA.
AC O9VZP8:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE CG12014 protein.
GN CG12014.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxId=7227;
RN [1]

```

RP SEQUENCE FROM N.A.
 RP STPAIN-Berkeley:
 RA MEDLINE=20196006; PubMed=10731132;
 RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abil U.F., Agbayan A., An H.-J., Andrews-Plannkch C., Baldwin D.,
 RA Ballew R.M., Baau A.A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Botkova D., Botchan M.R., Bouck J., Brockstein P., Brodtier P.,
 RA Buttis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cateley S., Dahlke C., Davenport L.B., Davies P.,
 RA De Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman I.J., Hernandez J.R., Houck D.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibbegan C.,
 RA Jatalai M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laiko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McInerney D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner C., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RL "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AEO03478; AAF47771.1; -
 DR FlyBase; PBgn0035445; CG12014.
 DR GO; GO:0008484; F:sulfuric ester hydrolase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR000917; Sulfatase.
 DR Pfam; PF00884; Sulfatase; 1.
 DR PROSITE; PS00523; SULFATASE 1; 1.
 SQ SEQUENCE 512 AA; 58904 MW; AA5B6D8400B6FB3A CRC64;

Query Match 48.9%; Score 44; DB 5; Length 512;
 Best Local Similarity 57.1%; Pred. No. 32;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0

1 VSKLEASTIREY 14
 : : : : :
 Db 278 ISPLQAQIRQSY 291

RESULT 4
 O8CRF1 PRELIMINARY; PRT; 116 AA.
 AC O8CRF1;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Conserved hypothetical protein.
 GN SEI840.
 OS *Staphylococcus epidermidis*.
 OC Bacteria; Firmicutes; Bacillales; *Staphylococcus*.
 OX NCBI_TaxID=1282;
 NP [1]
 NP SEQUENCE FROM N.A.

This Page Blank (usnto)

```
RESULT 14
ID UL61_HCMVA STANDARD; PRT; 431 AA.
AC P16818;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Hypothetical protein UL61.
GN UL61.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
SEQUENCE FROM N.A.
RA MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA Horsnell T., Hutchinson C.A. III, Kouzarides T., Martignetti J.A.,
RA Peddie B., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169."
RT Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17403; CA35376.1; -
DR F01; S09824; S09824.
KW Hypothetical protein.
SQ SEQUENCE 431 AA; 44309 MW; 61CC7288FA3B0743 CRC64;

Query Match 43.3%; Score 39; DB 1; Length 431;
Best Local Similarity 42.9%; Pred. No. 34;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 5 EASTRPOEYFGDA 18
Db 65 QATVRAEFTWGAA 78

RESULT 15
VF59_HUMAN STANDARD; PRT; 533 AA.
AC Q9HCL3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical zinc finger protein KIAA1559.
GN KIAA1559
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RA TISSUE=Brain;
RA MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hiroseawa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVIII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro."
RL DNA Res. 7:273-281(2000).
CC -1- FUNCTION: May function as a transcription factor.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -1- SIMILARITY: Contains 1 KRAB domain.
```

```
CC -----
CC -1- SIMILARITY: Contains 13 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB046779; BAB13385.1; ALT_INIT.
DR HSSP; P25490; IUBD.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF01352; KRAB; 1.
DR Pfam; PF00096; zf-C2H2; 13.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 13.
DR PROSITE; PS50805; KRAB; 1.
DR PROSITE; PS50028; ZINC_FINGER_C2H2_1; 13.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 13.
KW Hypothetical protein; Transcription regulation; DNA-binding;
KW Zinc-finger; Metal-binding; Nuclear protein; Repeat.
KW DOMAIN
FT ZN_FING 172 194 C2H2-TYPE 1.
FT ZN_FING 200 222 C2H2-TYPE 2.
FT ZN_FING 228 250 C2H2-TYPE 3.
FT ZN_FING 256 278 C2H2-TYPE 4.
FT ZN_FING 284 306 C2H2-TYPE 5.
FT ZN_FING 312 334 C2H2-TYPE 6.
FT ZN_FING 340 362 C2H2-TYPE 7.
FT ZN_FING 368 390 C2H2-TYPE 8.
FT ZN_FING 396 418 C2H2-TYPE 9.
FT ZN_FING 424 446 C2H2-TYPE 10.
FT ZN_FING 452 474 C2H2-TYPE 11.
FT ZN_FING 480 502 C2H2-TYPE 12.
FT ZN_FING 508 530 C2H2-TYPE 13.
SQ SEQUENCE 533 AA; 63463 MW; E40EF5EC22A99F10 CRC64;

Query Match 43.3%; Score 39; DB 1; Length 533;
Best Local Similarity 53.3%; Pred. No. 42;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 SKLEASTRPOEYFG 16
Db 127 SKIDGKEQDGGYFG 141
```

Search completed: September 10, 2004, 17:53:12
Job time : 9.24022 secs

DR PIR; A89944; A89944.
 DR HAMAP; MF_00035; -; 1.
 DR InterPro; IPR003156; DHHA1.
 DR InterPro; IPR002318; tRNA-synt_2c.
 DR InterPro; IPR006193; tRNA-synt_Ala.
 DR Pfam; PF02272; DHHA1; 1.
 DR Pfam; PF01411; tRNA-synt_2c; 1.
 DR PRINTS; PR00980; TRNASYNTHALA.
 DR TIGRFAMs; TIGR00344; alas; 1.
 DR PROSITE; PS50860; AA_TRNA_LIGASE_II_ALA; 1.
 DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 SQ SEQUENCE 876 AA; 98538 MW; 2B2BC79041AC264F CRC64;

Query Match 44.4%; Score 40; DB 1; Length 876;
 Best Local Similarity 61.5%; Pred. No. 48;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 VSKLEASTIROXY 13
 DB 1 MKKLKASEIRÖKY 13

RESULT 12
 SYA_STAAW STANDARD; PRT; 876 AA.
 AC Q8N67;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DE 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA ligase) (Alars).
 GN ALAS OR MM1568.
 OS Staphylococcus aureus (strain MM2).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 RX MEDLINE=196620;
 NCBI_TaxID=196620;
 RP SEQUENCE FROM N.A.
 RA BABA T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
 RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
 RA Yamamoto K., Hiramatsu K.;
 RA "Genome and virulence determinants of high virulence community-acquired MRSA";
 RT Lancet 359:1819-1827(2002).
 RL -1- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +
 CC diphosphate + L-alanyl-tRNA(Ala).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL; AF004827; BAB95433.1; -.
 DR HAMAP; MF_00035; -; 1.
 DR InterPro; IPR003156; DHHA1.
 DR InterPro; IPR002318; tRNA-synt_2c.
 DR InterPro; IPR006193; tRNA-synt_Ala.
 DR Pfam; PF02272; DHHA1; 1.
 DR Pfam; PF01411; tRNA-synt_2c; 1.
 DR PRINTS; PR00980; TRNASYNTHALA.
 DR TIGRFAMs; TIGR00344; alas; 1.
 DR PROSITE; PS50860; AA_TRNA_LIGASE_II_ALA; 1.
 DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 SQ SEQUENCE 876 AA; 98505 MW; 5D9D6628DDAEDFC CRC64;

Query Match 44.4%; Score 40; DB 1; Length 876;
 Best Local Similarity 61.5%; Pred. No. 48;

Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 VSKLEASTIROXY 13
 DB 1 MKKLKASEIRÖKY 13

RESULT 13
 PYRB_DEIRA STANDARD; PRT; 314 AA.
 AC Q9RVC0;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Aspartate carboxyltransferase (EC 2.1.3.2) (Aspartate
 DE transcarboxylase) (ATCase).
 GN PYRB OR DR1109.
 OS Deinococcus radiodurans.
 OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 RX MEDLINE=1239;
 NCBI_TaxID=1239;
 RP SEQUENCE FROM N.A.
 RA STRAIN=RI / ATCC 13939 / DSM 20539 / NCIB 9279;
 RX MEDLINE=2003696; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.D., Lam P., McDonald L., Uitterlinden T., Zalewski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RA "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans RI.";
 RL Science 286:1571-1577(1999).
 CC -1- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-aspartate = phosphate
 CC + N-carbamoyl-L-aspartate.
 CC -1- PATHWAY: Pyrimidine biosynthesis; second step.
 CC -1- SIMILARITY: Belongs to the ATCase/OTCase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL; AE001961; AAF10682.1; -.
 DR PIR; D75435; D75435.
 DR HSSP; P00479; 3CSU.
 DR TIGR; DR1109; -.
 DR HAMAP; MF_00001; -; 1.
 DR InterPro; IPR006130; Asp/Om Cotranf.
 DR InterPro; IPR002082; Asp carboxyltransf.
 DR InterPro; IPR006131; OTCase_O.
 DR InterPro; IPR006132; OTCase_P.
 DR Pfam; PF00185; OTCase; 1.
 DR Pfam; PF02729; OTCase_N; 1.
 DR PRINTS; PR00100; AOTCASE.
 DR TIGRFAMs; TIGR006670; asp carb tr; 1.
 DR PROSITE; PS00097; CARBAMOYLTRANSFERRASE; 1.
 DR Pyrimidine biosynthesis; Transferrase; Complete proteome.
 KW Pyrimidine biosynthesis; Transferrase; Complete proteome.
 SQ SEQUENCE 314 AA; 34081 MW; 631875A2AD3EF4B6 CRC64;

Query Match 43.3%; Score 39; DB 1; Length 314;
 Best Local Similarity 80.0%; Pred. No. 24;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 LEASTIROXY 13
 DB 148 LDAYTIRÖXY 157

DE Cystathionine beta-synthase (EC 4.2.1.22) (Serine sulphydrase)
 DE (Beta-ethionase).
 OS Cys4 OR STR4 OR YGR155W OR G6667.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Burkholderia; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycos.
 CC NCBI_Taxid=4932;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=X2180-1A;
 RC MEDLINE=93374830; PubMed=8366024;
 RA Charest H., Thomas D., Surdin-Kerjan Y.;
 RT "Cysteine biosynthesis in Saccharomyces cerevisiae occurs through the
 RT transsulfuration pathway which has been built up by enzyme
 RT recruitment.";
 RT J. Bacteriol. 175:5366-5374(1993).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=A5-8-1A.
 RC Ono B.I., Inoue T., Kijima K., Matsuda A., Negishi K., Shinoda S.;
 RA Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RC MEDLINE=96158062; PubMed=8585325;
 RA Skala U., Nawrocki A., Goffeau A.;
 RT "The sequence of a 27 kb segment on the right arm of chromosome VII
 RT from Saccharomyces cerevisiae reveals MOI1, NAT2, RPL30B, RSRI, CYS4,
 RT PEM1/CHO2, NSR1 genes and ten new open reading frames.";
 RT Yeast 11:1421-1427(1995).
 RN [4]
 RN SEQUENCE FROM N.A.
 RC STRAIN=S288c;
 RC MEDLINE=94294429; PubMed=8022826;
 RA Kruger W.D., Cox D.R.;
 RT "A yeast system for expression of human cystathionine beta-synthase:
 RT structural and functional conservation of the human and yeast genes.";
 RT Proc. Natl. Acad. Sci. U.S.A. 91:6614-6618(1994).
 CC -1- CATALYTIC ACTIVITY: L-serine + L-homocysteine = cystathionine +
 CC H(2)O.
 CC -1- COFACTOR: Pyridoxal phosphate.
 CC -1- PATHWAY: Homocysteine transsulfuration; first step.
 CC -1- SIMILARITY: Belongs to the cysteine synthase/cystathionine beta-
 CC synthase family.
 CC -1- SIMILARITY: Contains 1 CBS domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X72922; CAA51426.1; -;
 DR EMBL: D16502; BAA03952.1; -;
 DR EMBL: X85807; CAA59812.1; -;
 DR EMBL: Z72940; CAA97169.1; -;
 DR EMBL: L14578; AAC37401.1; -;
 DR PIR: A46661; A46661.
 DR HSSP: P35520; IJBO.
 DR GenBank: 141467; -;
 DR SGD: S0003387; CYS4.
 DR GO: GO:0005737; CysC; cytoplasm; IDA.
 DR GO: GO:0004122; P-cystathionine beta-synthase activity; IDA.
 DR GO: GO:0019344; P-cysteine biosynthesis; IMP.
 DR InterPro: IPR001926; B6_enzyme_beta.
 DR InterPro: IPR000644; CBS_domain.
 DR InterPro: IPR001216; Cys_synthase_BS.
 DR InterPro: IPR005857; Cysa_beta_synth.
 DR Pfam: PF00571; CBS; 2.
 DR Pfam: PF00291; PALP; 1.
 DR SMART: SM00116; CBS; 2.

DR TIGRfam: TIGR01137; cysa_beta; 1.
 DR PROSITE: P800901; CYS_SYNTHASE; 1.
 KW Cysteine biosynthesis; Lyase; Pyridoxal phosphate; CBS domain.
 FT BINDING 117 117
 FT DOMAIN 311 424
 FT CONFLICT 2 2
 FT CONFLICT 8 8
 FT CONFLICT 63 63
 FT CONFLICT 104 104
 FT CONFLICT 129 129
 FT CONFLICT 163 163
 FT CONFLICT 293 293
 FT CONFLICT 407 407
 FT CONFLICT 436 437
 FT CONFLICT 441 441
 FT CONFLICT 481 481
 FT CONFLICT 481 481
 SQ SEQUENCE 507 AA; 56021 MW; DDC7059B20FD0746 CRC64;
 Query Match 47.8%; Score 43; DB 1; Length 507;
 Best Local Similarity 58.8%; Pred. No. 7.8;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 2 SKLEASTROEYFGDA 18
 DB 351 SKLEASTTKYADVFGNA 367
 ID YN48 ARCFU STANDARD; PRT; 112 AA.
 AC 030321;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein AF2348.
 GN AF2348.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 CC NCBI_Taxid=22343;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RC MEDLINE=96049343; PubMed=9389475;
 RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kierlavage A.R., Graham D.E., Uitterback T.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirsnes E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uitterback T.,
 RA Cotton M.D., Spriggs T., Arllich P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AE001114; AAB91315.1; -;
 DR PIR: D69543; D69543.
 DR TIGR: AF2348; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 112 AA; 13012 MW; C9C6AC0ACD6AC730 CRC64;

```

RC STRAIN=FVB/N; TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.W., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stepleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Browneleim W.J., Usdin T.B., Toshiyuki S., Carrinon P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Raheyl J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boulfard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[3]
RN SEQUENCE OF 1-11 FROM N.A.
RP Groelz D., Bachmann M.,
RA Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: La protein plays a role in the transcription of RNA
CC polymerase III. It is most probably a transcription termination
CC factor. Binds to the 3' terminus of virtually all nascent
CC polymerase III transcripts. It is associated with precursor forms
CC of RNA polymerase III transcripts including rRNA and 4.5S, 5S, 7S,
CC and 7-2 RNAs.
CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: Phosphorylated (By similarity).
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L00993; AAA9415.1; -.
CC DR EMBL: BC003820; AAH03820.1; -.
CC DR EMBL: Y07951; CAA69249.1; -.
CC DR WGI: 98423; Ssb.
CC DR GO: 0005634; C:nucleus; IDA.
CC DR InterPro: IPR002344; Lupus_La.
CC DR InterPro: IPR006630; Lupus_La_dom.
CC DR InterPro: IPR005054; RNA_rec_mot.
CC DR Pfam: PF05383; La; 1.
CC DR Pfam: PF00076; rrm; 1.
CC DR PRINTS: PRO0302; LUPUSLA.
CC DR SMART: SM00715; LA; 1.
CC DR SMART: SMO0360; RRM; 1.
CC DR PROSITE: PS50102; RRM; 1.
CC DR PROSITE: PS00030; RRM_RNP_1; 1.
CC DR RNA-binding; Nuclear protein; Phosphorylation.
CC KW RNA-binding; Nuclear protein; Phosphorylation.
CC FT DOMAIN 111 187 RNA-BINDING (RRM).
CC SEQUENCE 415 AA; 47756 MW; 2D75197692FDC933 CRC64;
SQ

```

ID	LA_RAT	STANDARD	PRT	415 AA
AC	P3656:			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Lupus Ia protein homolog (Ia ribonucleoprotein) (Ia autoantigen homolog)			
DE	SSB OR SS-B.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93246255; PubMed=7916708;			
RA	Jensen I., Troester H., Bartsch H., Schwemmler M., Igloi G.L.,			
RA	Bachmann M.,			
RT	"Isolation of rat cDNA clones coding for the autoantigen SS-B/La:			
RL	detection of species-specific variations.",			
Gene	126:265-268(1993).			
CC	-1- FUNCTION: La protein plays a role in the transcription of RNA			
CC	polymerase III. It is most probably a transcription termination			
CC	factor. Binds to the 3' termini of virtually all nascent			
CC	polymerase III transcripts. It is associated with precursor forms			
CC	of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S,			
CC	and 7-2 RNAs.			
CC	-1- SUBUNIT: Interacts with DDX15 (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Nuclear (Probable).			
CC	-1- PTM: Phosphorylated (By similarity).			
CC	-1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.ebi.ac.uk/announcement/			
CC	or send an email to license@ebi.ac.uk).			
CC	-----			
DR	EMBL; X67859; CAA48043.1; -.			
DR	PIR; JCI494; JCI494.			
DR	InterPro; IPR002344; Lupus_La.			
DR	InterPro; IPR00630; Lupus_La_dom.			
DR	InterPro; IPR00504; RNA_rec_mot.			
DR	Pfam; PF05383; La; 1.			
DR	Pfam; PF00076; rrm; 1.			
DR	PRINTS; PR00302; LUPUSLA.			
DR	SMART; SM00715; LA; 1.			
DR	SMART; SM00360; RRM; 1.			
DR	PROSITE; PS0102; RRM; 1.			
DR	PROSITE; PS00030; RRM_RNP_1; 1.			
PT	RNA-binding; Nuclear protein; Phosphorylation.			
PT	DOMAIN 111 187 RNA-BINDING (RRM).			
SEQ	SEQUENCE 415 AA; 47777 MW; 033FD9CCE475F98 CRC64;			
Query Match	48.3%; Score 43.5; DB 1; Length 415;			
Best Local Similarity	55.6%; Pred. No. 5.1;			
Matches	10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;			
Oy	1 VSKLEASTIRQ-EYFFGD 17			
DB	10 MALEAKICHQIEYFGD 27			
RESULT 8				
ID	CBS_YEAST	STANDARD;	PRT;	507 AA.
AC	P32582; O05177;			
DT	01-OCT-1993 (Rel. 27, Created)			
DT	01-OCT-1993 (Rel. 27, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			

RP SEQUENCE FROM N.A.
 RC TISUB=Placenta, and Skeletal muscle;
 RA MEDLINE=22388257; PubMed=1477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stedlemon M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinot P., Prange C.,
 RA Rana S.S., Locuelli N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosnak S.A., McEwen K.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield V.S.N., Krzywinski M.T., Skalska U., Smallos D.E.,
 RA Scherch A., Schein J.E., Jones S.J.W., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [4]
 RP SEQUENCE OF 54-408 FROM N.A.
 RX MEDLINE=8819081; PubMed=2452201;
 RA Sturgees A.D., Peterson M.G., McNeillage L.J., Whittingham S.,
 RA Coppel R.S.;
 RT "Characteristics and epitope mapping of a cloned human autoantigen
 RT La.";
 RL J. Immunol. 140:3212-3218 (1988).
 RN [5]
 RP SEQUENCE OF 54-97 FROM N.A.
 RX MEDLINE=8516283; PubMed=3856888;
 RA Chambers J.C., Keene J.D.;
 RT "Isolation and analysis of cDNA clones expressing human lupus La
 RT antigen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2115-2119 (1985).
 RN [6]
 RP FUNCTION.
 RX MEDLINE=89251617; PubMed=2470590;
 RA Gottlieb E., Steitz J.A.;
 RT "Function of the mammalian La protein: evidence for its action in
 RT transcription termination by RNA polymerase III.";
 RL EMBO J. 8:851-861 (1989).
 RN [7]
 RP PHOSPHORYLATION.
 RX MEDLINE=97207017; PubMed=9054510;
 RA Fan H., Sakulich A.L., Goodier J.L., Zhang X., Qin J., Marale R.J.;
 RT "Phosphorylation of the human La antigen on serine 366 can regulate
 RT recycling of RNA polymerase III transcription complexes.";
 RL Cell 88:707-715 (1997).
 RN [8]
 RP INTERACTION WITH DDX15.
 RX MEDLINE=22346609; PubMed=12458796;
 RA Fournoux M.A., Kolkman M.J.M., Van der Heijden A., De Jong A.S.,
 RA Van Venrooij W.J., Pruijn G.J.M.;
 RT "The human La (SS-B) autoantigen interacts with DDX15/hPrp43, a
 RT putative DEAD-box RNA helicase.";
 RL RNA 8:1428-1443 (2002).
 RN [9]
 RP FUNCTION: La protein plays a role in the transcription of RNA
 RN polymerase III. It is most probably a transcription termination
 RN factor. Binds to the 3' termini of virtually all nascent
 RN polymerase III transcripts. It is associated with precursor forms
 RN of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S,
 RN and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE
 CC C-TERMINAL PART OF THE PROTEIN.
 CC -1- PTM: The N-terminus is blocked.
 CC -1- DISEASE: Sera from patients with systemic lupus erythematosus
 CC often contain antibodies that react with the normal cellular

CC La protein as if this antigen was foreign.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RNM) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation on
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-eb.ch/announce/>
 CC or send an email to license@isb-eb.ch).
 CC -----
 CC EMBL: X13697; CAA31985.1; -
 CC EMBL: J04205; AAA51885.1; -
 CC EMBL: BC001289; AAH01289.1; -
 CC EMBL: BC020818; AAH20818.1; -
 CC PIR: A11888; A11888.
 CC Genew: HGNC:11316; SSB.
 CC MIM: 109090; -
 DR GO: GO:0030529; C:ribonucleoprotein complex; TAS.
 DR GO: GO:0003729; F:RNA binding; TAS.
 DR GO: GO:0000049; F:RNA binding; TAS.
 DR GO: GO:0008334; P:histone mRNA metabolism; TAS.
 DR GO: GO:0006400; P:RNA modification; TAS.
 DR InterPro: IPR002344; Lupus_La.
 DR InterPro: IPR006630; Lupus_La_dom.
 DR InterPro: IPR00504; RNA_rec_mot.
 DR Pfam: PF05383; La; 1.
 DR Pfam: PF00076; rtm; 1.
 DR PRINTS: PR00302; LUPUSLA.
 DR SMART: SMO0715; LA; 1.
 DR SMART: SMO0360; RRM; 1.
 DR PROSITE: PS0102; RRM; 1.
 DR PROSITE: PS0030; RRM_RNP_1; 1.
 KW Systemic lupus erythematosus; RNA-binding; Phosphorylation;
 KW Nuclear protein.
 FT DOMAIN 111 187 RNA-BINDING (RRM).
 FT MOD RES 366 366 PHOSPHORYLATION (BY CK2).
 SQ SEQUENCE 408 AA; 46837 MW; EC153C15F9187FC4 CRC64;
 Query Match 48.3%; Score 43.5; DB 1; Length 408;
 Best Local Similarity 55.6%; Pred. No. 5;
 Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
 QY 1 VSKLFASTIRQ-EYFPGD 17
 Db 10 MALEAKICHQIEYFGD 27
 ID LA_MOUSE STANDARD; PRT; 415 AA.
 AC P32067;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
 DE homolog).
 OS SSB OR SS-B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OK NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93203630; PubMed=8454877;
 RA Topfer F., Gordon T., McCluskey J.;
 RT "Characterization of the mouse autoantigen La (SS-B). Identification
 RT of conserved RNA-binding motifs, a putative ATP binding site and
 RT reactivity of recombinant protein with poly(U) and human
 RL J. Immunol. 150:3091-3100 (1993).
 RN [2]
 RP SEQUENCE FROM N.A.

[1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=MAFR30309;
 RC MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asanizu E., Kato T., Sasamoto S.,
 RA Kaneko T., Ideasa Y., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shingo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.,
 RT Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 CC -1- CATALYTIC ACTIVITY: ATP + L-histidine + tRNA(His) = AMP +
 CC diophosphate + L-histidyl-tRNA(His).
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family..
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF003010; BAB53120.1; -.
 DR HAMAP; MF_00127; -; 1.
 DR InterPro; IPR004154; HGTP_anticonodon.
 DR InterPro; IPR004516; HisS.
 DR InterPro; IPR002314; tRNA-synt_2b.
 DR InterPro; IPR006195; tRNA_ligase_1t.
 DR Pfam; PF03129; HGTP_anticonodon; 1.
 DR Pfam; PF00587; tRNA-synt_2b; 1.
 DR TIGRFAMs; TIGR00442; hisS; 1.
 DR PROSITE; PS50862; AA tRNA LIGASE II; 1.
 KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 SQ SEQUENCE 500 AA; 54546 MW; 290E8B18CA2DF84 CRC64;
 Query Match 50.0%; Score 45; DB 1; Length 500;
 Best Local Similarity 55.6%; Pred. No. 3.4;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 QY 1 VKLEASTRQRYFGDA 18
 Db 398 VSELRAAGIRSEWYLGGA 415
 RESULT 4
 ID LA BOVIN STANDARD; PRT; 404 AA.
 AC P10881;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 26-FEB-2003 (Rel. 41, Last annotation update)
 DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
 DE homolog).
 DE SSB.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pituitary;
 RX MEDLINE=89202037; PubMed=2468131;
 RA Chan E.K.L., Sullivan K.F., Tan E.M.;
 RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus
 RT sequences for RNA-binding."
 RL Nucleic Acids Res. 17:2233-2244(1989).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA

polymerase III. It is most probably a transcription termination
 factor. Binds to the 3' termini of virtually all nascent
 polymerase III transcripts. It is associated with precursor forms
 of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S,
 and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DXI5 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE
 CC C-TERMINAL PART OF THE PROTEIN.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X13698; CAA31986.1; -.
 DR PIR; S03849; S03849.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 KM RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 111 187
 SQ SEQUENCE 404 AA; 46534 MW; 4EB30B5C262AD6A1 CRC64;
 Query Match 48.3%; Score 43.5; DB 1; Length 404;
 Best Local Similarity 55.6%; Pred. No. 5;
 Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
 QY 1 VKLEASTRQRYFGD 17
 Db 10 MAALRAKICHOIEYFGD 27
 RESULT 5
 ID LA HUMAN STANDARD; PRT; 408 AA.
 AC P05455;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein (Sjogren syndrome type B antigen) (SS-B) (La
 DE ribonucleoprotein) (La autoantigen).
 DE SSB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89202037; PubMed=2468131;
 RA Chan E.K.L., Sullivan K.F., Tan E.M.;
 RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus
 RT sequences for RNA-binding."
 RL Nucleic Acids Res. 17:2233-2244(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89053970; PubMed=3192525;
 RA Chambers J.C., Xenan D., Martin B.J., Keane J.D.;
 RT "Genomic structure and amino acid sequence domains of the human La
 RT autoantigen."
 RL J. Biol. Chem. 263:18043-18051(1988).
 RN [3]

DB 40 VSKLEASTIRQLEYFGDA 58

RESULT 2

LA DROME STANDARD; PRT; 390 AA.

AC P40796; Q24375; Q9VIN2;

DT 01-FEB-1995 (Rel. 31, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).

GN LA OR CG10922.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RP SEQUENCE FROM N.A.

RC STRAIN=Canton-S; TISSUE=Ovary;

RK MEDLINE=94309632; PubMed=8035794;

RA Bal C., Li Z., Tolias P.P.;

RT "Developmental characterization of a Drosophila RNA-binding protein homologous to the human systemic lupus erythematosus-associated La/SS-B autoantigen";

RT Mol. Cell. Biol. 14:5123-5129 (1994).

RN [2]

RP SEQUENCE FROM N.A.

RK MEDLINE=94309661; PubMed=8035818;

RA Yoo C.J., Molin S.L.;

RT "La proteins from Drosophila melanogaster and Saccharomyces cerevisiae: a yeast homolog of the La autoantigen is dispensable for growth";

RT Mol. Cell. Biol. 14:5412-5424 (1994).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RK MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle J.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle A.C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abail J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasly E.M., Beeson K.Y., Bencos P.V., Bertan B.P., Bhandari D., Bolshakov S., Borkov D., Botchan M.R., Bouck J., Brokstein P., Brothier P., Buttis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira J., Fleischmann W., Fowler C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J.R., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacile J.M., Palazuelos M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier B., Spradling A.C., Stapleton M., Strong R., Sun E., Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195 (2000).

CC -1- FUNCTION: May be involved in transcription termination by RNA polymerase III. Binds RNA and DNA. Binds to precursors of RNA polymerase III transcripts. May play a specialized role during fly development.

CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).

CC -1- DEVELOPMENTAL STAGE: Expressed throughout embryonic, larval, pupal, and adult development. Expression throughout the embryo is followed by a restricted pattern of mesodermal expression that is later confined to the visceral mesoderm, gonads, gut, and salivary glands.

CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.

CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.1sb-sib.ch/announce/> or send an email to license@1sb-sib.ch).

CC -----

DR EMBL; U07652; AAA20518.1; -;

DR EMBL; L32988; AAA21776.1; -;

DR EMBL; AE003666; AA53885.1; -;

DR PIR; A53773; A53773.

DR PIR; A53781; A53781.

DR FLYBase; FBgn0011638; La.

DR GO; GO:0008098; F:SS rRNA primary transcript binding; IDA.

DR GO; GO:0003723; F:RNA binding; NAS.

DR InterPro; IPR002344; Lupus_La.

DR InterPro; IPR006630; Lupus_La_dom.

DR InterPro; IPR00504; RNA_rec_mot.

DR Pfam; PF05383; La; 1.

DR Pfam; PF00076; rrm; 1.

DR PRINTS; PR00302; LUPUSLA.

DR SMART; SM00715; LA; 1.

DR SMART; SM00360; RRM; 1.

DR PROSITE; PS50102; RRM; 1.

DR PROSITE; PS00303; RRM_RNP_1; 1.

FT DOMAIN-binding; Nuclear protein; DNA-binding.

FT DOMAIN 149 234

FT CONFLICT 169 169 A -> T (IN REF. 1).

FT CONFLICT 182 183 KH -> NS (IN REF. 1).

FT CONFLICT 283 283 A -> R (IN REF. 1).

FT CONFLICT 329 329 K -> N (IN REF. 1).

FT SEQUENCE 390 AA; 44884 MW; A809928B89044655 CRC64;

SO Query Match 58.3%; Score 52.5; DB 1; Length 390;

Best Local Similarity 66.7%; Pred. No. 0.12;

Matches 12; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 2 SKLEASTIRQLEYFGDA 18

DB 48 TKQRAAIRQVEYFGDA 65

RESULT 3

SVL RHILQ STANDARD; PRT; 500 AA.

AC Q98770;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Histidyl-RNA synthetase (EC 6.1.1.21) (Histidine--tRNA ligase) (HISRS).

GN HIS OR MUR6928.

OS Rhizobium loti (Mesorhizobium loti).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Phyllobacteriaceae; Mesorhizobium.

OX NCBI_TaxID=381;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:01:41; Search time 7.24022 Seconds
(without alignments)
129,452 Million cell updates/sec

Title: US-09-836-073-18

Perfect score: 90
Sequence: 1 VSKLEASTTROEYFGDA 18

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79.5	88.3	383	1A_AEDAL	Q26457 aedes albop
2	52.5	58.3	390	1A_DROME	P40796 drosophila
3	45	50.0	500	1SYH_RHIL	O98760 rhizobium 1
4	43.5	48.3	404	1LA_BOVIN	P10981 bos taurus
5	43.5	48.3	408	1LA_HUMAN	P05455 homo sapien
6	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
7	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
8	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
9	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
10	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
11	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
12	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
13	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
14	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
15	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
16	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
17	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
18	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
19	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
20	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
21	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
22	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
23	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
24	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
25	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
26	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
27	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
28	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
29	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
30	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
31	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
32	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu
33	43.5	48.3	415	1LA_MOUSE	P32067 mus musculu

34	37.5	41.7	428	1	1A_XENLA	P28048 xenopus lae
35	37	41.1	268	1	1YAJ_HAETN	P44996 haemophilus
36	37	41.1	312	1	1KPS_CAUCR	Q9496 caulobacter
37	37	41.1	332	1	1MOD_DROME	P22816 drosophila
38	37	41.1	339	1	1MUG_THEMA	O9W774 thermocoga
39	37	41.1	422	1	1U83_CABEL	P34692 caenorhabdi
40	37	41.1	427	1	1AB_XENLA	P28049 xenopus lae
41	37	41.1	433	1	1SUCB_SCHPO	O94415 schizosacch
42	37	41.1	471	1	1YMT1_CABEL	P98080 caenorhabdi
43	37	41.1	550	1	1SYR_MYCSM	O9X500 mycobacteri
44	37	41.1	556	1	1SYR_LISIN	O92712 listeria in
45	37	41.1	556	1	1SYR_LISMO	O9Y493 listeria mo

ALIGNMENTS

RESULT 1	LA_AEDAL	STANDARD;	PRT;	383 AA.
AC	Q26457;			
DT	15-JUL-1998 (rel. 36, Last sequence update)			
DT	28-FEB-2003 (rel. 41, Last annotation update)			
DE	La protein homolog (la ribonucleoprotein) (la autoantigen homolog).			
OS	Aedes albopictus (Forest day mosquito).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.			
ON	NCBI_Taxid=1160;			
OX	(1)			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RX	MEDLINE=96135233; PubMed=8551578;			
RA	Pardigon N., Strauss J.H.;			
RT	"Mosquito homolog of the la autoantigen binds to Sindbis virus RNA.";			
RL	J. Virol. 70:1173-1181 (1996).			
CC	-1- FUNCTION: May be involved in transcription termination by RNA			
CC	polymerase III. Binds RNA and DNA. Binds to the 3' end of the			
CC	minus strand of Sindbis virus RNA. This may be significant for			
CC	Sindbis virus RNA replication.			
CC	-1- SUBCELLULAR LOCATION: Nuclear. Primarily nuclear, but significant			
CC	amounts are present in the cytoplasm.			
CC	-1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.			
CC	-1- SIMILARITY: TO VERTEBRATE PROTEIN 1A.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; S80954; AAB35931.1; -			
DR	InterPro; IPR002344; Lupus 1A.			
DR	InterPro; IPR006630; Lupus 1A dom.			
DR	InterPro; IPR000504; RNA_rec_mot.			
DR	Pfam; PF05383; la; 1.			
DR	Pfam; PF00076; trm; 1.			
DR	PRINTS; PR00715; LA; 1.			
DR	SMART; SM00715; LA; 1.			
DR	SMART; SM00360; RRM; 1.			
DR	PROSITE; PS50102; RRM; 1.			
DR	PROSITE; PS00030; RRM_RNP_1; FALSE NEG.			
KW	RNA-Binding; Nuclear protein; DNA-binding.			
KW	RNA-Binding; Nuclear protein; DNA-binding.			
FT	DOMAIN 141 228			
FT	SEQUENCE 383 AA; 44430 MW; 4E5CC8F21C40F452 CRC64;			
QY	1 VSKLEASTTROEYFGDA 18			
	Query Match	88.3%;	Score 79.5;	DB 1;
	Best Local Similarity	94.7%;	Pred. No. 1.8e-06;	Length 383;
	Matches 18;	Conservative 0;	Mismatches 0;	Indels 1;
				Gaps 1;

This Page Blank (uspto)

A:introns : 29/1

Query Match 45.6%; Score 41; DB 2; Length 633;
Best Local Similarity 53.3%; Pred. No. 37;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 KLEASTIROEYFGD 17
| : | | | : |
Db 353 KKDARTIROEYFLSD 367

RESULT 14

D70117

acridiavine resistance protein (acrB) homolog - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000
C:Accession: D70117
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kexlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
i Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; PMID:9805943; PMID:9403685
A:Accession: D70117
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1036 <KLE>
A:Cross-references: GB:AE000783; TIGR:BB0140
A:Experimental source: strain B31
C:Superfamily: cation efflux system membrane protein czcA

Query Match 45.6%; Score 41; DB 2; Length 1036;
Best Local Similarity 50.0%; Pred. No. 65;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 VSKLEASTIROEYFG 16
| | | | | : | | |
Db 82 VSKSESTIVSLRPHG 97

RESULT 15

D69543

hypothetical protein AF2348 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: D69543
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Ueterbach, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaime, B.P.; Sykes, S.
Smith, H.O.; Moese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; PMID:98049343; PMID:9389475
A:Accession: D69543
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-112 <KLE>
A:Cross-references: GB:AE001114; GB:AE000782; NID:92689437; PIDN:AAB91315.1; PID:9265074
C:Superfamily: Archaeoglobus fulgidus hypothetical protein AF2348

Query Match 44.4%; Score 40; DB 2; Length 112;
Best Local Similarity 47.1%; Pred. No. 7.6;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 SKLEASTIROEYFGDA 18
| : | | | : | | |
Db 3 SRSRKSVTLQDYRRGDS 19

Search completed: September 10, 2004, 18:02:39
Job time : 16.0782 secs

Query Match 46.7%; Score 42; DB 2; Length 451;
 Best Local Similarity 53.3%; Pred. No. 16;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 4 LEASTIROEYFGDA 18
 |||:|:|:|:|:|:|
 Db 187 LEANNAOKEYVLTDA 201

RESULT 9

T06586
 DNA-binding protein PD2 - garden pea
 C:Species: Pisum sativum (garden pea)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 08-Oct-1999
 C:Accession: T06586
 R:Sato, N.; Kazuno, A.A.; Ohta, N.; Ohshima, K.
 submitted to the EMBL Data Library, June 1996
 A:Description: Identification of a novel family of DNA-binding proteins with two AT-hook
 A:Reference number: Z15774
 A:Accession: T06586
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-632 <SAT>
 A:Cross-references: EMBL:X98740; NID:e995229; PIDN:CAA67292.1; PID:e275185
 A:Experimental source: cv. Alaska

Query Match 46.7%; Score 42; DB 2; Length 632;
 Best Local Similarity 50.0%; Pred. No. 24;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 LEASTIROEYFGD 17
 |||:|:|:|:|:|:|
 Db 161 VEATVSDDEYTYGD 174

RESULT 10

H83743
 ABC transporter (permease) BH0752 [imported] - Bacillus halodurans (strain C-125)
 C:Species: Bacillus halodurans
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: H83743
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Mieno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hirai
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: H83743
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-642 <STO>
 A:Cross-references: GB:AP001509; GB:BA000004; NID:g10173176; PIDN:BAE04471.1; GSPDB:GN00
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH0752

Query Match 46.7%; Score 42; DB 2; Length 642;
 Best Local Similarity 53.8%; Pred. No. 24;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 EASTIROEYFGD 17
 |||:|:|:|:|:|:|
 Db 472 EKALVHBEYFGD 484

RESULT 11

C96636
 hypothetical protein T7P1.19 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: C96636
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: C96636
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-322 <STO>
 A:Cross-references: GB:AE005173; NID:g6751694; PIDN:AAF27677.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: T7P1.19
 A:Map position: 1

Query Match 45.6%; Score 41; DB 2; Length 322;
 Best Local Similarity 64.3%; Pred. No. 17;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VSKLEASTIROEY 14
 |||:|:|:|:|:|:|
 Db 44 VSKLMSQIRRPY 57

RESULT 12

B87591
 cytochrome c-type biogenesis protein Cych [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 30-Sep-2002
 C:Accession: B87591
 R:Nierman, W.C.; Feldblym, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.U.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolton,
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of Caulobacter crescentus.
 A:Reference number: A87249; MUID:21173698; PMID:11259647
 A:Accession: B87591
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-370 <STO>
 A:Cross-references: GB:AE005673; NID:g13424358; PIDN:AAK24726.1; GSPDB:GN00148
 C:Genetics:
 A:Gene: CC2762
 C:Superfamily: cych protein

Query Match 45.6%; Score 41; DB 2; Length 370;
 Best Local Similarity 43.8%; Pred. No. 20;
 Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 KLEASTIROEYFGDA 18
 |||:|:|:|:|:|:|
 Db 220 KVDAADVARYVYLGRA 235

RESULT 13

T06703
 hypothetical protein T29H11.90 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 22-Oct-1999
 C:Accession: T06703
 R:Querier, F.; Choisme, N.; Robert, C.; Brottier, P.; Wincker, P.; Catolico, L.; Artigau,
 submitted to the Protein Sequence Database, April 1999
 A:Reference number: Z15793
 A:Accession: T06703
 A:Molecule type: DNA
 A:Residues: 1-633 <QUE>
 A:Cross-references: EMBL:AL049659; GSPDB:GN00061; ATSP:T29H11.90
 A:Experimental source: cultivar Columbia; BAC clone T29H11
 C:Genetics:
 A:Gene: ATSP:T29H11.90
 A:Map position: 3

Best Local Similarity 66.7%; Pred. No. 0.17;
Matches 12; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Oy 2 SKLEASTIRQ-EYFGDA 18
Db 48 TKQERAIRIROYEYFGDA 65

RESULT 3

T10213
hypothetical protein F25G13.200 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999

C:Accession: T10213

R:Chan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Meyer, K.F.X.; Lemcke

submitted to the Protein Sequence Database, June 1999

A:Reference number: Z16991

A:Accession: T10213

A:Molecule type: DNA

A:Residues: 1-365 <BEV>

A:Cross-references: EMBL:X13697; NID:936414; PIDN:CAA31985.1; PID:936415

A:Experimental source: cultivar Columbia; BAC clone F25G13

C:Genetics:

A:Gene: ATSP:F25G13.200

A:Map position: 4

A:Introns: 55/2

Query Match 48.9%; Score 44; DB 2; Length 365;
Best Local Similarity 38.9%; Pred. No. 5.5;
Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Oy 1 VSKLEASTIROEYFGDA 18
Db 8 ILRCSSCTWREYFPHDA 25

RESULT 4

S03849
ribonucleoprotein la - bovine

N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B

C:Species: Bos primigenius taurus (cattle)

C:Date: 31-Dec-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C:Accession: S03849

R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.

Nucleic Acids Res. 17, 2233-2244, 1989

A:Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences

A:Reference number: S03848; MUID:89202037; PMID:2468131

A:Accession: S03849

A:Molecule type: mRNA

A:Residues: 1-404 <CHA>

A:Cross-references: EMBL:X13698; NID:9755; PIDN:CAA31986.1; PID:9756

A>Note: part of this sequence was confirmed by protein sequencing

C:Comment: This protein associates with a variety of small RNA molecules, most of which

act as a transcription termination factor.

C:Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology

C:Keywords: blocked amino end; phosphoprotein; RNA binding

F:113-118/Domain: ribonucleoprotein repeat homology <RNM>

F:113-118/Region: RNA-binding RNP2 motif

F:151-158/Region: RNA-binding RNP1 motif

F:128-404/Domain: phosphorylated #status predicted <PHY>

Query Match 48.3%; Score 43.5; DB 1; Length 404;
Best Local Similarity 55.6%; Pred. No. 7.6;
Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

Oy 1 VSKLEASTIRQ-EYFGD 17
Db 10 MAALAKICHQIEYFGD 27

RESULT 5

A31888
ribonucleoprotein la - human

N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B; Sjogren syndrome antigen

C:Species: Homo sapiens (man)

C:Date: 21-May-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C:Accession: A31888; S03848; A22956; A61051; S11013; I55553; I70205; I70206; A31273

R:Chambers, J.C.; Kenan, D.; Martin, B.J.; Keene, J.D.

J. Biol. Chem. 263, 18043-18051, 1988

A:Title: Genomic structure and amino acid sequence domains of the human La autoantigen.

A:Reference number: A31888; MUID:89053970; PMID:3192525

A:Accession: A31888

A:Molecule type: mRNA

A:Residues: 1-408 <CHA>

A:Cross-references: GB:J04205; NID:9178686; PIDN:AAA51885.1; PID:9178687

R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.

Nucleic Acids Res. 17, 2233-2244, 1989

A:Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences

A:Reference number: S03848; MUID:89202037; PMID:2468131

A:Accession: S03848

A:Molecule type: mRNA

A:Residues: 1-408 <CH2>

A:Cross-references: EMBL:X13697; NID:936414; PIDN:CAA31985.1; PID:936415

R:Chambers, J.C.; Keene, J.D.

Proc. Natl. Acad. Sci. U.S.A. 82, 2115-2119, 1985

A:Title: Isolation and analysis of cDNA clones expressing human lupus La antigen.

A:Reference number: A22956; MUID:85166283; PMID:3856888

A:Accession: A22956

A:Molecule type: mRNA

A:Residues: 45-97; 'LK' <CH3>

A:Cross-references: GB:J04205

A>Note: this sequence has been revised in reference A31888

R:Nyman, U.; Ringertz, N.R.; Petersson, I.

Immunol. Lett. 22, 65-72, 1989

A:Title: Demonstration of an amino terminal La epitope recognized by human anti-La sera.

A:Reference number: A61051; MUID:89379261; PMID:2476379

A:Accession: A61051

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-19; 'E', '21-47 <NYM>

R:Sturgess, A.D.; Peterson, M.G.

J. Immunol. 140, 3212-3218, 1988

A:Title: Characteristics and epitope mapping of a cloned human autoantigen La.

A:Reference number: S11013; MUID:88199081; PMID:2425201

A:Accession: S11013

A:Molecule type: mRNA

A:Residues: 'E', '55-287', 'V', '289-408 <STU>

A:Cross-references: EMBL:M20328; NID:9337456; PIDN:AAA6577.1; PID:9337457

R:Kohsaka, H.; Yamamoto, K.; Fujii, H.; Mura, H.; Miyasaka, N.; Nishioaka, K.; Miyamoto, J.

Clin. Invest. 85, 1566-1574, 1990

A:Title: Fine epitope mapping the human SS-B/La protein: Identification of a distinct au

A:Reference number: I55553; MUID:90237237; PMID:1692037

A:Accession: I55553

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 81-107 <RES>

A:Cross-references: GB:M5261; NID:9338491; PIDN:AAA36652.1; PID:9338495

A:Accession: I70205

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 174-224 <RE2>

A:Cross-references: GB:M5263; NID:9338492; PIDN:AAA36653.1; PID:9338496

A:Accession: I70206

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 279-342 <RE3>

A:Cross-references: GB:M5262; NID:9338493; PIDN:AAA36654.1; PID:9338497

C:Comment: This protein associates with a variety of small RNA molecules, most of which

act as a transcription termination factor.

C:Genetics:

A:Gene: GDB:SSB

A:Cross-references: GDB:125359; OMIM:109090

A:Map position: 2

A:Introns: 22/3; 57/2; 115/3; 151/3; 185/2; 209/2; 223/3; 264/3; 380/2

C:Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology

C:Keywords: phosphoprotein; RNA binding

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:33:11; Search time 14.0782 Seconds
(without alignments)
122.988 Million cell updates/sec

Title: US-09-836-073-18

Perfect score: 90
Sequence: 1 VSKLEASTIRQRYFGDA 18

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR 78:*

1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52.5	58.3	390	2 A53773	La/SS-B homolog D-
2	52.5	58.3	390	2 A53781	ribonucleoprotein
3	44	48.9	365	2 T10213	hypothetical prote
4	43.5	48.3	404	1 S03849	ribonucleoprotein
5	43.5	48.3	408	1 A31888	ribonucleoprotein
6	43.5	48.3	415	1 JCI194	ribonucleoprotein
7	43	47.8	507	2 A48661	cystathionine beta
8	42	46.7	451	2 AB2296	UDP-N-acetylglucos
9	42	46.7	632	2 T06586	DNA-binding protei
10	42	46.7	642	2 H83743	ABC transporter (p
11	41	45.6	322	2 C96636	hypothetical prote
12	41	45.6	370	2 B87591	cytochrome c-type
13	41	45.6	633	2 D0703	hypothetical prote
14	41	45.6	1036	2 D70117	acetylcholine resist
15	40	44.4	112	2 D69543	hypothetical prote
16	40	44.4	379	2 A95130	glycogen biosynthe
17	40	44.4	379	2 G98000	required for glyco
18	40	44.4	546	2 S29522	casein kinase I ho
19	40	44.4	835	2 T30030	hypothetical prote
20	40	44.4	876	2 A89944	alanyl-tRNA synthe
21	40	44.4	1201	2 H86434	protein F17F8.21 f
22	39.5	43.9	506	2 H64618	sigma-54 interacti
23	39	43.3	149	2 T08274	hypothetical prote
24	39	43.3	258	2 E71646	hypothetical prote
25	39	43.3	279	2 E97863	hypothetical prote
26	39	43.3	314	2 D75435	aspartate carbamoy
27	39	43.3	426	2 T06086	hypothetical prote
28	36	43.3	431	2 T21594	hypothetical prote
29	39	43.3	431	2 S09824	hypothetical prote

30	39	43.3	517	2 S12227	hypothetical prote
31	39	43.3	550	1 A4936	arginine-tRNA ligase
32	39	43.3	801	2 T07617	proteinase TMP - t
33	39	43.3	866	2 E72113	Clp proteinase ATP
34	39	43.3	866	2 H86508	Clp proteinase ATP
35	39	43.3	872	2 H81556	ATP-dependent Clp
36	39	43.3	904	2 AF2304	poly(A) polymerase
37	39	43.3	942	2 S75598	DNA-binding protei
38	38.5	42.8	211	2 G83718	hypothetical prote
39	38.5	42.8	215	2 A69787	probable DNA bindi
40	38.5	42.8	215	2 AG1704	a probable DNA bin
41	38.5	42.8	215	2 AH1333	complement factor
42	38.5	42.8	1234	1 NEMSH	hypothetical prote
43	38	42.2	173	2 T45501	hypothetical prote
44	38	42.2	202	2 T15874	hypothetical prote
45	38	42.2	212	2 C85706	probable outer mem

ALIGNMENTS

```

RESULT 1
A53773
La/SS-B homolog D-1a - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #ext_change 21-Jul-2000
C/Accession: A53773
R/Bai, C.; Li, Z.; Tolias, P.P.
Mol. Cell. Biol. 14, 5123-5129, 1994
A/Title: Developmental characterization of a Drosophila RNA-binding protein homologous
A/Reference number: A53773; UID:94309632; PMID:8035794
A/Accession: A53773
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-390 <BAI>
A/Cross-references: GB:U07652; NID:9464019; PIDN:AAA20518.1; PID:9464020
C/Genetics:
A/Gene: FlyBase:la
A/Cross-references: FlyBase:FBgn0011638
C/Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology
C/Keywords: leucine zipper; RNA binding

Query Match          58.3%; Score 52.5; DB 2; Length 390;
Best Local Similarity 66.7%; Pred. No. 0.17;
Matches 12; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY      2 SKLEASTIRQ-EYFGDA 18
Db      48 TKQERAIHQVEYFGDA 65

RESULT 2
A53781
ribonucleoprotein La - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #ext_change 20-Sep-1999
C/Accession: A53781
R/Yoo, C.J.; Wolin, S.L.
Mol. Cell. Biol. 14, 5412-5424, 1994
A/Title: La proteins from Drosophila melanogaster and Saccharomyces cerevisiae: a yeast
A/Reference number: A53781; UID:94309661; PMID:8035818
A/Accession: A53781
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-390 <YOO>
A/Cross-references: GB:L32988; NID:9488469; PID:9488470
C/Genetics:
A/Gene: FlyBase:la
A/Cross-references: FlyBase:FBgn0011638
C/Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology
C/Keywords: RNA binding

Query Match          58.3%; Score 52.5; DB 2; Length 390;

```


; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO: 19
 ; LENGTH: 16
 ; TYPE: PRT
 ; ORGANISM: Drosophila
 US-09-836-073-19

Query Match 52.6%; Score 50; DB 9; Length 16;
 Best Local Similarity 57.1%; Pred. No. 0.37;
 Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
 QY 4 DQRIKQLEYFGN 17
 Db 2 ERAIRQVEIYFGD 15

Search completed: September 10, 2004, 18:11:57
 Job time : 41.424 secs

```

RESULT 11
US-10-424-599-254664
; Sequence 254664, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424.599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 254664
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_71985C.1.pep
US-10-424-599-254664

```

Query Match	53.7%	Score 51;	DB 12;	Length 395;
Best Local Similarity	52.9%	Pred. No. 7.6;		
Matches	9;	Conservative	4;	Mismatches 4;
				Indels 0;
				Gaps 0.
Q7	1	DDADDRITKQLEYYFGN	17	
	:		:	
Db	101	EDLKRITKIQVEYYPSD	117	

```

RESULT 12
US-10-767-701-46290
: Sequence 46290, Application US/10767701
: Publication NO. US20040172684A1
: GENERAL INFORMATION:
: APPLICANT: Kovalic, David K.
: APPLICANT: Zhou, Yihua
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
: FILE REFERENCE: 38-21(513535)B
: CURRENT APPLICATION NUMBER: US/10/767,701
: CURRENT FILING DATE: 2004-01-29
: NUMBER OF SEQ ID NOS: 63128
: SEQ ID NO 46290
: LENGTH: 489
: TYPE: PRT
: ORGANISM: Sorghum bicolor
: FEATURE:
: OTHER INFORMATION: Clone ID: SORBI-28MAY03-C18446_1.pep
US-10-767-701-46290

```

Query Match	53.7%	Score 51;	DB 16;	Length 489;
Best Local Similarity	50.0%	Pred. No. 9.5;		
Matches	9;	Conservative	4;	Mismatches 5;
				Indels 0;
				Gaps 0;

```
Qy      1 DDADQRIKQLEYFFGNI 18
         | : | | | | | : |
Db      133 DEVVHKITQVEYFFSDI 150
```

```

RESULT 43
US-10-425-114-69520
; Sequence 69520, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E

```

```

: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53313)B
: CURRENT APPLICATION NUMBER: US/10/425,114
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 73128
: SEQ ID NO 69520
: LENGTH: 490
: TYPE: PRT
: ORGANISM: Zea mays
: FEATURE:
: OTHER INFORMATION: Clone ID: UC-ZMFLB73030C03_F11.pep
US-10-425-114-69520

```

Query Match	53.7%	Score 51	DB 12	Length 490
Best Local Similarity	50.0%	Pred. No. 9.5		
Matches	9	Conservative	4	Mismatches 5; Indels 0; Gaps 0.
QY	1	DDADQRIIKOLEYTFGN	18	
	:	: :	:	:
Db	131	DEVNKKTKQVEYFSDI	148	

```

RESULT 14
US-10-437-963-106767
; Sequence 106767, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 106767
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_11181C.1.pep
US-10-437-963-106767

```

Query Match	53.7%	Score 51;	DB 16;	length 511;
Best Local Similarity	50.0%	Pred. No. 10;		
Matches	9;	Conservative	4;	Mismatches 5;
				Indels 0;
				Gaps 0;

QY 1 DDADQRIKQLEYFNGNI 18
| : : | | : | | : |
Db 117 DEVVHKITQVEYFYFSDI 134

RESULT 15
US-09-836-073-19
Sequence 19, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Das, S.
APPLICANT: Balda, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 220002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630

Best Local Similarity 52.9%; Pred. No. 2.4;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 DDADRIIKOLEYFGN 17
|:|:|:|:|:|:
Db 132 DELCDRIYKQVEYFSD 148

RESULT 7

US-10-437-963-179489
; Sequence 179489, Application US/10437963
; Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Mu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 179489

LENGTH: 453

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_76947C.1.pep

US-10-437-963-179489

Query Match
Best Local Similarity 56.8%; Score 54; DB 16; Length 453;
Pred. No. 2.9;

Matches 7; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 DDADRIIKOLEYFGN 18
|:|:|:|:|:|:
Db 105 EDVARRVKQVEYFSDV 122

US-10-437-963-134637
; Sequence 134637, Application US/10437963
; Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Mu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 134637

LENGTH: 405

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_36391C.1.pep

US-10-437-963-134637

Query Match
Best Local Similarity 55.8%; Score 53; DB 16; Length 405;
Pred. No. 3.7;

Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 DDADRIIKOLEYFGN 17
|:|:|:|:|:|:
Db 85 DELCDRIYKQVEYFSD 101

RESULT 9

US-10-424-599-254661

; Sequence 254661, Application US/10424599
; Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei

APPLICANT: Zhou Yihua
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 254661

LENGTH: 143

TYPE: PRT

ORGANISM: Glycine max

FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_71982C.1.pep

US-10-424-599-254661

Query Match
Best Local Similarity 53.7%; Score 51; DB 12; Length 143;
Pred. No. 2.6;

Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 DDADRIIKOLEYFGN 17
|:|:|:|:|:|:
Db 101 EDLKLKIKQVEYFSD 117

US-10-425-114-66921
; Sequence 66921, Application US/10425114
; Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei

APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53313)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 66921

LENGTH: 357

TYPE: PRT

ORGANISM: Zea mays

FEATURE:
OTHER INFORMATION: Clone ID: UC-ZMFLB71148E12_Flt1.pep

US-10-425-114-66921

Query Match
Best Local Similarity 53.7%; Score 51; DB 12; Length 357;
Pred. No. 6.8;

Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADRIIKOLEYFGN 18
|:|:|:|:|:|:
Db 163 DDPVKIKQVEYFSDI 180

```

; APPLICANT: Zhou Yihua
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 258877
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_75790C.1.pep
US-10-424-599-258877

Query Match      62.1%; Score 59; DB 12; Length 174;
Best Local Similarity 50.0%; Pred. No. 0.17;
Matches          9; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY              1 DDADQRITKOLEYYFGNI 18
                  |||::|||::|||::
Db              156 DEASOKILNQVEYYFSDL 173

RESULT 3
US-10-424-599-258874
; Sequence 258874, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 258874
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(467)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_75788C.1.pep
US-10-424-599-258874

Query Match      62.1%; Score 59; DB 12; Length 467;
Best Local Similarity 50.0%; Pred. No. 0.48;
Matches          9; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY              1 DDADQRITKOLEYYFGNI 18
                  |||::|||::|||::
Db              148 DEASOKILNQVEYYFSDL 165

RESULT 4
US-10-177-478-8
; Sequence 8, Application US/10177478
; Publication No. US20030106590A1
; GENERAL INFORMATION:
; APPLICANT: Dang, Van-Dinh
; APPLICANT: Okamuro, Jack
; TITLE OF INVENTION: Chimeric Histone Acetyltransferase
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 11696-037001
; CURRENT APPLICATION NUMBER: US/10/177,478
; CURRENT FILING DATE: 2002-06-21

```

```

; PRIOR APPLICATION NUMBER: U.S. 60/300,135
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-177-478-8

Query Match          60.0%   Score 57; DB 14; Length 411;
Best Local Similarity 52.9%; Pred. No. 0.87;
Matches 9; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY      1 DDADRRIRKOLEYYFGN 17
       |:|::||::|||::|:
DB      90 DELNQRIKROYEYFSD 106

RESULT 5
US-09-836-073-15
; Sequence 15, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Xenopus
US-09-836-073-15

Query Match          57.9%   Score 55; DB 9; Length 18;
Best Local Similarity 56.2%; Pred. No. 0.067;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      2 DADQRIRIKOLEYYFGN 17
       ||::|::|||::|:
DB      2 DLDTKICEQIEYFPGD 17

RESULT 6
US-10-767-701-45524
; Sequence 45524, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5355)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 45524
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C33606_1.pep
US-10-767-701-45524

Query Match          56.8%   Score 54; DB 16; Length 376;

```


ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-26

Query Match 48.4%; Score 46; DB 3; Length 39;
Best Local Similarity 63.6%; Pred. No. 0.63; 0; Indels 0; Gaps 0;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 IIKQLEYYFGN 17
|:|:|:|:|:|:
DB 1 ILRQVEYYFGD 11

RESULT 15
US-09-430-323-26
Sequence 26, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin
Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-430-323-26

Query Match 48.4%; Score 46; DB 4; Length 39;
Best Local Similarity 63.6%; Pred. No. 0.63; 0; Indels 0; Gaps 0;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 IIKQLEYYFGN 17
|:|:|:|:|:|:
DB 1 ILRQVEYYFGD 11

Search completed: September 10, 2004, 18:05:11
Job time: 16.8883 secs

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-26

Query Match 48.4%; Score 46; DB 3; Length 39;
Best Local Similarity 63.6%; Pred. No. 0.63;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 IIRQVEYFGN 17
|:::|||||:
Db 1 IIRQVEYFGD 11

RESULT 13
US-08-974-549A-216
Sequence 216, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morlin, Gregg B.
APPLICANT: Harley, Calvin H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 216:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-216

Query Match 48.4%; Score 46; DB 3; Length 39;
Best Local Similarity 63.6%; Pred. No. 0.63;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 IIRQVEYFGN 17
|:::|||||:
Db 1 IIRQVEYFGD 11

RESULT 14
US-08-854-050-26
Sequence 26, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morlin, Gregg B.
APPLICANT: Harley, Calvin H.
TITLE OF INVENTION: No. 6261836el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESSES:

Query Match	50.5%	Score 48	DB 5	Length 385
Best Local Similarity	50.0%	Pred. No. 3.9'		
Matches	9	Conservative	4	Mismatches 5; Indels 0; Gaps 0
Qy	1	DDADQRIIKQLEYYRGN	18	
Db	256	DQSDQVILIKLNIHVGNI	273	

RESULT 8
US-09-187-859-41

```

? Patent No 6358920
?
? GENERAL INFORMATION:
?
? APPLICANT: Blaschuk, Orest W.
?
? APPLICANT: Gout, Barbara J.
?
? TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
?
? TITLE OF INVENTION: CADERLIN-MEDIATED FUNCTIONS
?
? FILE REFERENCE: 100086.40701
?
? CURRENT APPLICATION NUMBER: US/09/187,859A
?
? CURRENT FILING DATE: 1998-11-06
?
? NUMBER OF SEQ ID NOS: 4052
?
? SOFTWARE: PatentIn Ver. 2.0
?
? SEQ ID NO 41
?
? LENGTH: 108
?
? TYPE: PR1
?
? ORGANISM: Mus musculus
?
US-09-187-859-41

```

	Query Match	49.5%	Score 47,	DB 4,	Length 108;
	Best Local Similarity	47.1%;	Pred. No. 1.4;		
	Matches	8; Conservative	5; Mismatches	4; Indels	0;
Qy	2 DADQRIRIKOLEYFGNT	18			
	: : :				
Db	26 DADGINKKILYFFSNL	42			

```

RESULT 9
US-09-839-542B-41
: Sequence 41. Application US/09839542B
: Patent No. 656996
: GENERAL INFORMATION:
: APPLICANT: Blaschuk, Orest W.
: APPLICANT: Symonch, James Matthew
: TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
: FILE REFERENCE: 100086,40701
: CURRENT APPLICATION NUMBER: US/09/839,542B
: CURRENT FILING DATE: 2001-04-20
: NUMBER OF SEQ. ID NOS: 4052
: SOFTWARE: PatentIn Ver. 2.0
: SEQ. ID NO. 41
: LENGTH: 108
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-839-542B-41

```

Query Match	49.5%	Score 47	DB 4	Length 108
Best Local Similarity	47.1%	Pred. No. 1.4		
Matches	8	Conservative	5	Mismatches 4
				Indels 0
				Gaps 0
QY	2	DADQRIRKQLEYFGNI	18	
Db	26	DADGEINKELLYFNSL	42	

RESULT 10
US-09-535-852-41
; Sequence 41, Application US/09535852
; Patent No. 6638911
; GENERAL INFORMATION:

```

      APPLICANT: Blachuk, Orest W.
      APPLICANT: Symonds, James M.
      APPLICANT: Gour, Barbara J.
      TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
      TITLE OF INVENTION: DESMOSOMAL CADHERIN-MEDIATED FUNCTIONS
      FILE REFERENCE: 100086.407C6
      CURRENT APPLICATION NUMBER: US/09/535,852
      CURRENT FILING DATE: 2001-05-21
      NUMBER OF SEQ ID NOS: 2009
      SOFTWARE: FastSeq for Windows Version 4.0
      SEQ ID NO 41
      LENGTH: 108
      TYPE: PRT
      ORGANISM: Mus musculus
      US-09-535-852-41

Query Match          49.5%; Score 47; DB 4; Length 108;
Best Local Similarity 47.1%; Pred. No. 1.4;
Matches      8; Conservative      5; Mismatches      0; Gaps      0;

QY      2 DADQRIIKOLEYVFGNI 18
      |||:|::|:|:|
Db      26 DADGINKELIYFNSNL 42

```

```

RESULT 11
US-09-540-236-2229
; Sequence 2229, Application US/09540236
; Patent No. 6673810
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540.236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2229
; LENGTH: 191
; TYPE: PRT
; ORGANISM: M.catarhalis
US-09-540-236-2229

```

Query Match	49.5%	Score 47	DB 4	Length 191
Best Local Similarity	47.1%	Pred. No. 2.6		
Matches	8	Conservative	5	Mismatches
			4	Indels
			0	Gaps
			0	
QY	1	DDADQRIIKQLEHYFFGN	17	
		::: :::		
Db	140	DDLEPQLIKLEPFHSH	156	

RESULT 12
US-08-851-843A-26
: Sequence 26, Application US/08851843A
: Patent No. 6093809
: GENERAL INFORMATION:
: APPLICANT: Cecch, Thomas R,
: APPLICANT: Lingner, Joachim
: APPLICANT: Nakamura, Toru
: APPLICANT: Chapman, Karen B.
: APPLICANT: Morlin, Gregg B.
: APPLICANT: Harley, Calvin
: APPLICANT: Andrews, William H.
: TITLE OF INVENTION: NO. 6093809e1 Telomerase
: NUMBER OF SEQUENCES: 225
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, 8th Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: United States of America
: Zip: 94111

REFERENCE/DOCKET NUMBER: 44301-Z/JPW/AGL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-516-801-2

Query Match 50.5%; Score 48; DB 2; Length 385;
Best Local Similarity 50.0%; Pred. No. 3.9;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIIKOLEYFYFNGI 18
|:||||:|:|:|
Db 256 DQSDQVITIKLNIHVNGI 273

RESULT 5
US-08-248-355-2
Sequence 2, Application US/08248355
PATENT INFORMATION:
APPLICANT: Goff, Stephen P.
TITLE OF INVENTION: A CDNA Clone of the Human Ini-1 Gene Encoding
TITLE OF INVENTION: A Protein That Binds to the HIV-1 Integrase;
TITLE OF INVENTION: and Its Use in Antiviral Therapy
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/248,355
FILING DATE: 24-May-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: White Esq., John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 44301
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-248-355-2

Query Match 50.5%; Score 48; DB 3; Length 385;
Best Local Similarity 50.0%; Pred. No. 3.9;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIIKOLEYFYFNGI 18
|:||||:|:|:|
Db 256 DQSDQVITIKLNIHVNGI 273

RESULT 6

US-09-167-206-16
Sequence 16, Application US/09167206A
PATENT No. 6476193
GENERAL INFORMATION:
APPLICANT: Nandabalan, Krishnan
APPLICANT: Schulz, Vincent P.
APPLICANT: Yang, MeiJa
TITLE OF INVENTION: NIK1 PROTEIN AND NIK1 PROTEIN COMPLEXES
FILE REFERENCE: 15966-521 NIK1 protein complexes
CURRENT APPLICATION NUMBER: US/09/167,206A
CURRENT FILING DATE: 1998-10-06
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 385
TYPE: PRT
ORGANISM: Homo sapiens
US-09-167-206-16

Query Match 50.5%; Score 48; DB 4; Length 385;
Best Local Similarity 50.0%; Pred. No. 3.9;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIIKOLEYFYFNGI 18
|:||||:|:|:|
Db 256 DQSDQVITIKLNIHVNGI 273

RESULT 7
PCT-US95-06683-2
Sequence 2, Application PC/TUS9506683
GENERAL INFORMATION:
APPLICANT: Goff, Stephen P.
TITLE OF INVENTION: A CDNA Clone of the Human Ini-1 Gene Encoding
TITLE OF INVENTION: A Protein That Binds to the HIV-1 Integrase;
TITLE OF INVENTION: and Its Use in Antiviral Therapy
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06683
FILING DATE: 24-May-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/248,355
FILING DATE: 24-May-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: White Esq., John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 44301-A-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-06683-2

Db 77 DOSDQVITIKINHVGNI 94

RESULT 2

US-08-248-355-3
Sequence 3, Application US/08248355

GENERAL INFORMATION:

APPLICANT: Goff, Stephen P.
APPLICANT: Kalpana, Ganjam V.

TITLE OF INVENTION: A cDNA Clone of the Human Ini-1 Gene Encoding
TITLE OF INVENTION: a Protein That Binds to the HIV-1 Integrase;

TITLE OF INVENTION: and Its Use in Antiviral Therapy

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York

STATE: New York
COUNTRY: U.S.A.

ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/248,355
FILING DATE: 24-May-1994

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: White Esq., John P.
REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 44301

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 204 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: N

ANTI-SENSE: N
FRAGMENT TYPE: N-terminal

US-08-248-355-3

Query Match 50.5%; Score 48; DB 3; Length 204;
Best Local Similarity 50.0%; Pred. No. 1.9;

Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DDADQRIIKOLEYFYFNGI 18
Db 77 DOSDQVITIKINHVGNI 94

RESULT 3
PCT-US95-06683-3

Sequence 3, Application PC/TUS9506683

GENERAL INFORMATION:

APPLICANT: Goff, Stephen P.
APPLICANT: Kalpana, Ganjam V.

TITLE OF INVENTION: A cDNA Clone of the Human Ini-1 Gene Encoding
TITLE OF INVENTION: a Protein That Binds to the HIV-1 Integrase;

TITLE OF INVENTION: and Its Use in Antiviral Therapy

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York

STATE: New York
COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/06683
FILING DATE: 24-May-1995

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/248,355
FILING DATE: 24-May-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: White Esq., John P.
REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 44301-A-PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 391-0525
TELEFAX: (212) 278-0400

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 204 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: N

ANTI-SENSE: N
FRAGMENT TYPE: N-terminal

PCT-US95-06683-3

Query Match 50.5%; Score 48; DB 5; Length 204;
Best Local Similarity 50.0%; Pred. No. 1.9;

Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DDADQRIIKOLEYFYFNGI 18
Db 77 DOSDQVITIKINHVGNI 94

RESULT 4
US-08-516-801-2

Sequence 2, Application US/08516801

Patent No. 5872213

GENERAL INFORMATION:

APPLICANT: Goff, Stephen P.
APPLICANT: Kalpana, Ganjam V.

TITLE OF INVENTION: A cDNA Clone of the Human Ini-1 Gene Encoding a
TITLE OF INVENTION: Protein That Binds to the HIV-1 Integrase; and Its Use in Anti

TITLE OF INVENTION: and Its Use in Antiviral Therapy

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York

STATE: New York
COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/516,801
FILING DATE: 18-August-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: John P. White, Esq.
REGISTRATION NUMBER: 28,678

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:43:01; Search time 15.8893 Seconds
(without alignments)
58.488 Million cell updates/sec

Title: US-09-836-073-17

Perfect score: 95

Sequence: 1 DDADQRIIKQLEVFNGNI 18

Scoring table:

Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B COMB.pep:*
3: /cgn2_6/prodata/2/iaa/6A COMB.pep:*
4: /cgn2_6/prodata/2/iaa/6B COMB.pep:*
5: /cgn2_6/prodata/2/iaa/6C COMB.pep:*
6: /cgn2_6/prodata/2/iaa/6D COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	48	50.5	204	US-08-516-801-3
2	48	50.5	204	US-08-248-355-3
3	48	50.5	204	PCT-US95-06683-3
4	48	50.5	385	US-08-516-801-2
5	48	50.5	385	US-08-248-355-2
6	48	50.5	385	US-09-167-206-16
7	48	50.5	385	PCT-US95-06683-2
8	47	49.5	108	US-09-187-859-41
9	47	49.5	108	US-09-839-542B-41
10	47	49.5	108	US-09-535-852-41
11	47	49.5	191	US-09-540-236-2229
12	46	48.4	39	US-08-851-843A-26
13	46	48.4	39	US-08-974-549A-216
14	46	48.4	39	US-08-854-050-26
15	46	48.4	39	US-09-430-323-26
16	46	48.4	39	US-09-402-181B-216
17	46	48.4	39	US-09-721-456-216
18	45	47.4	1076	US-09-470-443-6
19	45	47.4	1145	US-09-470-443-2
20	45	47.4	1145	US-09-470-443-4
21	45	47.4	3074	US-09-543-681A-5508
22	44	46.3	18	US-09-316-630-3
23	44	46.3	18	US-09-316-630-4
24	43	45.3	38	US-08-851-843A-25
25	43	45.3	38	US-08-974-549A-215
26	43	45.3	38	US-08-854-050-25
27	43	45.3	38	US-09-430-323-25

28	43	45.3	38	US-09-402-181B-215	Sequence 215, App
29	43	45.3	38	US-09-721-456-215	Sequence 215, App
30	43	45.3	498	US-09-107-532A-6991	Sequence 6991, App
31	41	43.2	21	US-08-475-955-20	Sequence 20, App
32	41	43.2	108	US-09-187-855-43	Sequence 43, App
33	41	43.2	108	US-09-839-542B-43	Sequence 43, App
34	41	43.2	108	US-09-535-852-43	Sequence 43, App
35	41	43.2	265	US-09-328-352-4347	Sequence 4347, App
36	41	43.2	626	US-09-489-039A-9106	Sequence 9106, App
37	41	43.2	775	US-08-966-388-4	Sequence 4, App
38	41	43.2	775	US-09-188-403-4	Sequence 4, App
39	41	43.2	775	US-09-188-404-4	Sequence 4, App
40	41	43.2	775	US-09-281-259-4	Sequence 4, App
41	40	42.1	108	US-09-149-476-552	Sequence 552, App
42	40	42.1	185	US-08-984-295-1	Sequence 1, App
43	40	42.1	186	US-09-149-476-194	Sequence 394, App
44	40	42.1	335	US-09-540-236-2030	Sequence 2030, App
45	39	41.1	72	US-09-107-532A-3822	Sequence 3822, App

ALIGNMENTS

RESULT 1
US-08-516-801-3
Sequence 3, Application US/08516801
Patent No. 5872213
GENERAL INFORMATION:
APPLICANT: Gott, Stephen P.
APPLICANT: Kalpana, Ganjam V.
TITLE OF INVENTION: A CDNA Clone of the Human Int-1 Gene Encoding a
TITLE OF INVENTION: Protein That Binds to the HIV-1 Integrase; and Its Use in Anti-
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSES: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/516,801
FILING DATE: 18-August-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: John P. White, Esq.
REGISTRATION NUMBER: 28,578
REFERENCE/DOCKET NUMBER: 44301-Z/UPW/AGL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526
INFORMATION FOR SEQ. ID NO. 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 204 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: N
ANTI-SENSE: N
FRAGMENT TYPE: N-terminal
US-08-516-801-3
Query Match 50.5%; Score 48; DB 2; Length 204;
Best Local Similarity 50.0%; Pred. No. 1.9; Indels 0; Gaps 0;
Matches 9; Conservative 4; Mismatches 5

DB 51 ERAIROYEYFGD 64

RESULT 10

ID Q940X9 PRELIMINARY; PRT; 826 AA.

AC Q940X9;

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE AT5G21160/T10F18_190 (Proline-rich protein family).

GN AT5G21160.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; euroids II; Brassicales; Brassicaceae; Arabidopsis.

OC NCB1_TaxID=3702;

OX [1]

RP SEQUENCE FROM N.A.

RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C., Bann J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.J., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S., Shiozaki K., Davis R.W., Theologis A., Ecker J.R.;

RT Arabidopsis cDNA clones."

RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Shinn P., Chen H., Cheuk R., Kim C.J., Bowser L., Carninci P., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y., Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Palm C.J., Quach H.J., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C., Yamada K., Yu G., Yuan S., Shiozaki K., Davis R.W., Theologis A., Ecker J.R.;

RT Arabidopsis ORF clones."

RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA STRAIN-CV. Columbia;

RA Town C.D., Haas B.J., Maiti R., Hannick L.I., Chan A.P., Rongning C.M., Smith Jr R.K., Arabogast T., Tallon L.V., Utterback T.R., Vanhaken S.E., Feldblum T.V., Yu C., Wortman J.R., White O., Fraser C.M.;

RT "Arabidopsis thaliana chromosome 5 BAC F13M11 genomic sequence."

RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY052365; AAK96556.1; -

DR EMBL: AY139801; AAM98107.1; -

DR EMBL: AC140977; AAO73903.1; -

DR GO: GO:0003743; P:translational initiation factor activity; IEA.

DR GO: GO:0006413; P:translational initiation; IEA.

DR InterPro: IPR006630; Lupus la dom.

DR InterPro: IPR002965; P rich extensn.

DR InterPro: IPR001950; TIF_SUTL.

DR Pfam: PF05383; Ia; 1.

DR PRINTS: PR01217; PRICHEXTENS.

DR PROSITE: PS01118; SUTL 1; 1.

SO SEQUENCE 826 AA; 91377 MW; 8D4192285B609D9A CRC64;

QY Query Match 52.6%; Score 50; DB 10; Length 826; Best Local Similarity 50.0%; Pred. No. 18; Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

DB 2 DADRIIKOLEYFGN 17

277 DLRDVKOYEFSD 292

RESULT 11

Q77881

ID Q77881 PRELIMINARY; PRT; 93 AA.

AC Q77881;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE Envelope protein (Fragment).

GN ENV.

OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OC NCB1_TaxID=11676;

OX [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=95191002; PubMed=7884875;

RA Mulder-Kampinga G.A., Simonon A., Kuiken C.L., Dekker J., Scherpbier H.J., de Perre P., Boer K., Goudemits J.;

RT "Similarity in env and gag genes between genomic RNAs of human immunodeficiency virus type 1 (HIV-1) from mother and infant is unrelated to time of HIV-1 RNA positivity in the child."

RL J. Virol. 69:2285-2296(1995).

DR EMBL: Z47894; CAA87908.1; -

DR GO: GO:0019028; C:viral capsid; IEA.

DR GO: GO:0019031; C:viral envelope; IEA.

DR GO: GO:0005198; P:structural molecule activity; IEA.

DR InterPro: IPR000777; GP120.

DR Pfam: PF00516; GP120; 1.

KW AIDS; Coat protein; Glycoprotein.

FT NON_TER 1 93

FT NON_TER 1 93

SO SEQUENCE 93 AA; 10561 MW; 4941876B9B6FFA8 CRC64;

QY Query Match 51.6%; Score 49; DB 15; Length 93; Best Local Similarity 61.5%; Pred. No. 2.9; Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DB 5 QRIIKOLEYFGN 17

73 QKVTLKOEYFGN 85

RESULT 12

ID Q91WT4 PRELIMINARY; PRT; 170 AA.

AC Q91WT4;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE Env glycoprotein (Fragment).

GN ENV.

OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OC NCB1_TaxID=11676;

OX [1]

RP SEQUENCE FROM N.A.

RA STRAIN=85CD225;

RA MEDLINE=21134754; PubMed=11242522;

RA Yang C., Dash B., Hanna S.L., Frances H.S., Nzilambi N., Colquhoun R.C., St Louis M., Quinn T.C., Polke T.M., Lal R.B.;

RT "Prevalence of HIV type 1 subtype G among commercial sex workers from Kinshasa, Democratic Republic of Congo."

RL AIDS Res. Hum. Retroviruses 17:361-365(2001).

DR EMBL: AF260442; AAF71590.1; -

DR GO: GO:0019028; C:viral capsid; IEA.

DR GO: GO:0019031; C:viral envelope; IEA.

DR GO: GO:0005198; P:structural molecule activity; IEA.

DR InterPro: IPR000777; GP120.

DR Pfam: PF00516; GP120; 1.

KW AIDS; Coat protein; Glycoprotein.

FT NON_TER 1 170

FT NON_TER 1 170

SO SEQUENCE 170 AA; 19351 MW; 177B41DAF6F7B7E CRC64;

QY Query Match 51.6%; Score 49; DB 15; Length 170; Best Local Similarity 52.9%; Pred. No. 5.3;


```
RT clone:01414.E05";
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AP003375; BAB90138.1; -.
DR Gramene; Q85078; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR002344; LupaLs_La_dom.
DR InterPro; IPR006630; LupaLs_La_dom.
DR Pfam; PF05383; La; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
SQ SEQUENCE 389 AA; 42100 MW; 10C1863EAB6FA7B7 CRC64;

Query Match 55.8%; Score 53; DB 10; Length 389;
Best Local Similarity 52.9%; Pred. No. 2.7;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 DDADRIKOLEYFGN 17
Db 85 DELCDRIYKQVYFSD 101

RESULT 7
ID Q80H15 PRELIMINARY; PRT; 206 AA.
AC Q80H15.
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Untranslated region binding-protein.
GN UBP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauromia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA L'Ecuyer T.J., Fang H.-L.;
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF467897; AAL76269.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR002344; LupaLs_La_dom.
DR InterPro; IPR006630; LupaLs_La_dom.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 206 AA; 23992 MW; 965B62F7DFPB90E9 CRC64;

Query Match 54.7%; Score 52; DB 13; Length 206;
Best Local Similarity 57.1%; Pred. No. 2.1;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 4 DORIKOLEYFGN 17
Db 14 ESKICQLEIYFGN 27

RESULT 8
ID Q7ZBB0 PRELIMINARY; PRT; 211 AA.
AC Q7ZBB0.
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
```

```
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TM08A7146;
RX MEDLINE=22483018; PubMed=12596722;
RA Kapiga S., Basse M.;
RA Kijelu I.E., Remylo B., Chaplin B., Sam N., Nkya W.M.M., Shao J.,
RT "HIV Type 1 Subtypes Among Bar and Hotel Workers in Moshi, Tanzania.";
RL AIDS Res. Hum. Retroviruses 19:57-64 (2003).
DR EMBL; AY123866; AAM95850.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR007777; GP120.
DR Pfam; PF00516; GP120; 1.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 211 AA; 23512 MW; D94732D39FC1497B CRC64;

Query Match 52.6%; Score 50; DB 15; Length 211;
Best Local Similarity 69.2%; Pred. No. 4.5;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 5 ORIKOLEYFGN 17
Db 71 QVAKOLEIYFGN 83

RESULT 9
ID Q8TBV5 PRELIMINARY; PRT; 390 AA.
AC Q8TBV5.
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE AT2034P.
GN LA OR CG10922.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY075257; AAL68124.1; -.
DR FLYbase; FBGN011638; LA.
DR GO; GO:0008088; F:5S rRNA primary transcript binding; IDA.
DR GO; GO:0003723; F:RNA binding; NAS.
DR InterPro; IPR002344; LupaLs_La_dom.
DR InterPro; IPR006630; LupaLs_La_dom.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 390 AA; 44869 MW; 797FDE26B903C909 CRC64;

Query Match 52.6%; Score 50; DB 5; Length 390;
Best Local Similarity 57.1%; Pred. No. 8.5;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 4 DORIKOLEYFGN 17
Db 14 ESKICQLEIYFGN 27
```

```

DR GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR005054; RNA_rec_moc.
DR Pfam; PF00076; itm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50102; RRM; 1.
SQ SEQUENCE 427 AA; 4896 MW; 1E7CD82D8A9C69A CRC64;

Query Match
Best Local Similarity 57.9%; Score 55; DB 13; Length 427;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 DDADRIIKQLEYFGN 17
DB 11 DDTRICEQIEYFGD 26

RESULT 5
080567 PRELIMINARY; PRT; 545 AA.
AC 080567; Q9CSX1;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Expressed protein (VirF-interacting protein FIP1) (Hypothetical
protein) (At2g43970/F6E13.10)
GN AT2G43970 OR F6E13.10/AT2G43970.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_Taxid=3702;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Columbia;
RC Roundley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L.,
RA Brandon R.C., Sykes S.M., Mason T.M., Kertavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Mayda E., Tzfira T., Citovsky V.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,
RA Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,
RA Koeseema B., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shim P.,
RA Southwick A., Tracy S.E., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Koeseema B., Meyers M.C., Shin P.,
RA Banh J., Bowser L., Carninci P., Chung M.K., Goldsmith A.D.,

```

```

RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yu G., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K.,
RA Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Banh J., Bowser L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [9]
RP SEQUENCE FROM N.A.
AC EMBL; AC004005; AAC23405.2; -
DR EMBL; AF332565; AAK06847.1; -
DR EMBL; AY056238; AAL07087.1; -
DR EMBL; AF375410; AAK52994.1; -
DR EMBL; AF367277; AAK56266.1; -
DR EMBL; AY129474; AAM91060.1; -
DR PIR; T00677; T00677.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR Pfam; PF05383; La; 1.
DR PRINTS; PR00302; LUPUSLA.
KW Hypothetical protein.
SQ SEQUENCE 545 AA; 60589 MW; E1A933261FE1ED80 CRC64;

Query Match
Best Local Similarity 56.8%; Score 54; DB 10; Length 545;
Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 DDADRIIKQLEYFGN 18
DB 191 EDStQKIVQVYFSD 208

RESULT 6
080578 PRELIMINARY; PRT; 389 AA.
AC 080578;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE RNA-binding protein-like.
GN Orl414_R05.3.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_Taxid=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Banh J., Bowser L., Carninci P., Chung M.K., Goldsmith A.D.,

```

```

OS Arabidopsis thaliana (Mouse-ear cress) .
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta .
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids .
OX NCBT_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shun P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Bahr J., Bowler L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamita A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones";
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shun P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
RA Hewan J.W., Ishida J., Jones T., Kamita A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis ORF clones";
RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN EMBL, AY050403; AA91419.1; -.
DR EMBL, BT000588; AAA18157.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006300; Lupus_La_dom.
DR InterPro; IPR005054; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR Pfam; PF0076; Itm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PSS10102; RRM; 1.
SQ SEQUENCE 422 AA; 46842 MW; 4ECABBF1E068F0E CRC64;

Query Match          60.0%; Score 57; DB 10; Length 422;
Best Local Similarity 52.9%; Pred. No. 0.65;
Matches      9; Conservative    6; Mismatches     2; Indels    0; Gaps    0;

Qy      1 DDADORTIKOLEYFYFGN 17
       |:|::|||::|||:
Db      101 DELNQKIIROVEYFSD 117

RESULT 4
O7ZTK2 PRELIMINARY; PRT; 427 AA.
AC O7ZTK2;
DT 01-JUN-2003 (TrEMBLrel_24, Created)
DT 01-JUN-2003 (TrEMBLrel_24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel_25, Last annotation update)
DE Similar to lupus la protein homolog B.
OS Xenopus laevis (African clawed frog) .
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBT_TaxId=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Klein S., Strauberg R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
RN EMBL, BC046554; AAH46554.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:29:01; Search time 41.229 Seconds
(without alignments)
137.751 Million cell updates/sec

Title: US-09-836-073-17

Perfect score: 95

Sequence: 1 DDADQRIKQLEFYFGNT 18

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL.25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriophage.*
- 17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	96.8	396	5	001806
2	57	60.0	411	10	09FL36
3	57	60.0	422	10	094A38
4	55	57.9	427	13	072TK2
5	54	56.8	545	10	080567
6	53	55.8	389	10	0850T8
7	52	54.7	206	13	080H15
8	50	52.6	211	15	07ZB80
9	50	52.6	390	5	08TRV5
10	50	52.6	826	10	0940X9
11	49	51.6	93	15	077881
12	49	51.6	170	15	091WT4
13	48	50.5	90	15	069711
14	48	50.5	341	15	07Z0S0
15	48	50.5	342	15	07Z0R9
16	48	50.5	342	15	07Z0P5

17	48	50.5	342	15	07Z0P4	Q7ZQP4 human immun
18	48	50.5	342	15	07Z0P3	Q7ZQP3 human immun
19	48	50.5	376	11	09ESU0	Q9ESU0 mus musculus
20	48	50.5	385	4	09UBH2	Q9UBH2 homo sapien
21	48	50.5	385	4	09ESU1	Q9ESU1 mus musculus
22	48	50.5	394	4	09H836	Q9H836 homo sapien
23	48	50.5	397	16	08NOL9	Q8NOL9 corynebacte
24	48	50.5	859	15	011947	Q11947 human immun
25	48	50.5	863	15	08JBB4	Q8JBB4 human immun
26	48	50.5	890	4	08N4R3	Q8N4R3 homo sapien
27	47	49.5	102	15	09Q0A7	Q9Q0A7 human immun
28	47	49.5	114	15	09YL33	Q9YL33 human immun
29	47	49.5	115	15	09YL52	Q9YL52 human immun
30	47	49.5	115	15	074722	Q74722 human immun
31	47	49.5	143	15	09J6A7	Q9J6A7 human immun
32	47	49.5	343	10	094LD0	Q94LD0 oryza sativ
33	47	49.5	401	13	07ZT10	Q7ZT10 brachydantio
34	47	49.5	898	11	088193	Q88193 mus musculus
35	47	49.5	934	11	091Y15	Q91Y15 mus musculus
36	46	48.4	102	15	090AW6	Q90AW6 human immun
37	46	48.4	102	15	090AN7	Q90AN7 human immun
38	46	48.4	102	15	090AY3	Q90AY3 human immun
39	46	48.4	128	15	09QMB7	Q9QMB7 human immun
40	46	48.4	342	15	07Z0R1	Q7Z0R1 human immun
41	46	48.4	351	15	07Z0Q3	Q7Z0Q3 human immun
42	46	48.4	351	15	07Z0Q2	Q7Z0Q2 human immun
43	46	48.4	351	15	07Z0Q1	Q7Z0Q1 human immun
44	46	48.4	351	15	07Z0Q0	Q7Z0Q0 human immun
45	46	48.4	351	15	07Z0P9	Q7ZQP9 human immun

ALIGNMENTS

RESULT 1

ID 001806 PRELIMINARY; PRT; 396 AA.

AC 001806; (T-EMBLrel. 04, Created)

DT 01-JUL-1997 (T-EMBLrel. 04, Last sequence update)

DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)

DE C4AE4.4 protein.

GN C4AE4.4.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Pelodermidae; Caenorhabditis.

OC Rhabditidae; Pelodermidae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RX MEDLINE=99069613; PubMed=9851916;

RA None;

RT "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium."

RL Science 282:2012-2018(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RT "The sequence of C. elegans cosmid C4AE4.";

RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RT Submitted (JUG-1999) to the EMBL/GenBank/DBJ databases.

RL EMBL; AF003140; AAB54169.1; -

DR PIR; T30953; T30953.

DR WormPep; C4AE4.4; CE08718.

DR GO; GO:0005634; C:nucleus; IEA.

DR GO; GO:0003723; F:RNA binding; IEA.

DR InterPro; IPR002344; Lupus La.

DR InterPro; IPR006630; Lupus_La_dom.

```

DR EMBL, X67859; CAA48043.1; -
DR PIR, JCI494; JCI494.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La.dom.
DR InterPro; IPR000504; RNA_rec_mol.
DR Pfam; PF05383; La; 1.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS00102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
DR Rfam-binding; Nuclear Protein; Phosphorylation.
FT DOMAIN 111 187
SQ SEQUENCE 415 AA; 47777 MW; 033FD9CCE1475F98 CRC64;

Query Match
Best Local Similarity 46.3%; Score 44; DB 1; Length 415;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 DORIRIKOLEYFGN 17
Db 14 EAKICHOIEYFGD 27

RESULT 15
DPOE MOUSE STANDARD; PRT; 2283 AA.
AC 09WVF7; Q9OX50;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE DNA polymerase epsilon, catalytic subunit A (EC 2.7.7.7) (DNA
DE polymerase II subunit A).
GN POLE OR POLE1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP MEDLINE=99296371; PubMed=10366722;
RX Huang D., Knutti R., Palosaeari H., Pospiech H., Sytaoja J.E.;
RT "cDNA and structural organization of the gene Pole1 for the mouse DNA
RT polymerase epsilon catalytic subunit."
RL Biochim. Biophys. Acta 1445:363-371(1999).
CC -1- FUNCTION: Participates in DNA repair and in chromosomal DNA
CC replication.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + (DNA) (N).
CC -1- SUBUNIT: Consists of two subunits (258 kDa and 55 kDa).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DOMAIN: The DNA polymerase activity domain resides in the
CC N-terminal half of the protein, while the C-terminus is necessary
CC for complexing subunits B and C. The C-terminus may also regulate
CC the catalytic activities of the enzyme.
CC -1- SIMILARITY: Belongs to the DNA polymerase type-B family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL, AF126382; AAD46482.1; JOINED.
DR EMBL, AF126383; AAD46482.1; JOINED.
DR EMBL, AF126384; AAD46482.1; JOINED.
DR EMBL, AF126385; AAD46482.1; JOINED.
DR EMBL, AF126386; AAD46482.1; JOINED.
DR EMBL, AF126387; AAD46482.1; JOINED.
DR EMBL, AF126388; AAD46482.1; JOINED.
DR EMBL, AF126389; AAD46482.1; JOINED.
DR EMBL, AF126390; AAD46482.1; JOINED.
DR EMBL, AF126391; AAD46482.1; JOINED.
DR EMBL, AF126392; AAD46482.1; JOINED.
DR EMBL, AF126393; AAD46482.1; JOINED.
DR EMBL, AF126394; AAD46482.1; JOINED.
DR EMBL, AF126395; AAD46482.1; JOINED.
DR EMBL, AF126396; AAD46482.1; JOINED.
DR EMBL, AF126397; AAD46482.1; JOINED.
DR MCD; MGI:1196391; Pole.
DR InterPro; IPR006172; DNA_pol_B.
DR InterPro; IPR006134; DNA_pol_B_dom.
DR InterPro; IPR006133; DNA_pol_B_exo.
DR Pfam; PF00136; DNA_pol_B; 1.
DR SMART; SM00486; POLBC; 1.
DR Transferrase; DNA-directed DNA polymerase; DNA replication; DNA repair;
KW DNA-binding; Zinc-finger; Nuclear protein.
FT ZN FING 2155 2235
FT ZN FING 205 205
FT ZN FING 284 284
FT ZN FING 1309 1309
FT ZN FING 1320 1320
FT ZN FING 1386 1388
FT ZN FING 1665 1665
FT ZN FING 2086 2086
FT ZN FING 2111 2111
FT ZN FING 2283 AA; 262025 MW; 7005C34354E465F3 CRC64;
SQ SEQUENCE 2283 AA; 262025 MW; 7005C34354E465F3 CRC64;

Query Match
Best Local Similarity 46.3%; Score 44; DB 1; Length 2283;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 DADORIRIKOLEYFG 15
Db 73 DEDKRLVSAVDYDF 86

Search completed: September 10, 2004, 17:53:10
Job time : 8.24022 secs

```

Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 DORIKOLEYFGN 17
: : : : :
Db 14 EAKICHQIEYFGD 27

RESULT 13

LA_MOUSE STANDARD; PRT; 415 AA.

AC P32067;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
GN SSB OR SS-B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=93203630; PubMed=8454877;
RA Topfer F., Gordon T., McCluskey J.;
RT "Characterization of the mouse autoantigen La (SS-B). Identification of conserved RNA-binding motifs, a putative ATP binding site and reactivity of recombinant protein with poly(U) and human autoantibodies.";
RT J. Immunol. 150:3091-3100(1993).
RN 12

SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Brownstein M.J., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E., Raha S.S., Loughell N.A., Peters G.J., Abramson R.D., Mullen S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywicki M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
RN 13

SEQUENCE OF 1-11 FROM N.A.
RP Groetz D., Bachmann M.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' termin of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S, and 7-2 RNAs.
CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: Phosphorylated (By similarity).
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; L00993; AAA39415.1; -
CC EMBL; BC003820; AA03820.1; -
CC EMBL; Y07951; CA69249.1; -
CC MGD; MG1:98423; Ssb.
CC GO; GO:0005634; C:nucleus; IDA.
CC InterPro; IPR002344; Lupus_La.
CC InterPro; IPR006630; Lupus_La.
CC InterPro; IPR00504; RNA_rec_mot.
CC Pfam; PF00076; rrm; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00715; LA; 1.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS50102; RRM; 1.
CC PROSITE; PS00030; RRM_RNP_1; 1.
CC RNA-binding; Nuclear protein; Phosphorylation.
CC DOMAIN 111 187 RNA-BINDING (RRM).
SQ SEQUENCE 415 AA; 47756 MW; 2D75197692FDC933 CRC64;

Query Match 46.3%; Score 44; DB 1; Length 415;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 DORIKOLEYFGN 17
: : : : :
Db 14 EAKICHQIEYFGD 27

RESULT 14

LA_RAT STANDARD; PRT; 415 AA.

AC P38656;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
GN SSB OR SS-B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=93246255; PubMed=7916708;
RA Semsei I., Troester H., Bartsch H., Schwemmler M., Igloi G.L., Bachmann M.;
RT "Isolation of rat cDNA clones coding for the autoantigen SS-B/La: detection of species-specific variations.";
RT Gene 126:265-268(1993).
CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' termin of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S, and 7-2 RNAs.
CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: Phosphorylated (By similarity).
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DB 14 EAKICHQIEYVFGD 27

RESULT 12

LA_HUMAN STANDARD; PRT; 408 AA.

AC P05455;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-JUL-1989 (Rel. 11, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Lupus La protein (Sjogren syndrome type B antigen) (SS-B) (La ribonucleoprotein) (La autoantigen).

DE SSB.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A. PubMed=2468131;

RX MEDLINE=892020377; PubMed=2468131;

RA Chan E.K.L., Sullivan K.F., Tan E.M.;

RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences for RNA-binding."

RL Nucleic Acids Res. 17:2233-2244 (1989).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=89053970; PubMed=3192525;

RA Chambers J.C., Kenan D., Martin B.J., Keene J.D.;

RT "Genomic structure and amino acid sequence domains of the human La autoantigen."

RL J. Biol. Chem. 263:18043-18051 (1988).

RN [3]

RP SEQUENCE FROM N.A.

RX TISSUE=Placenta, and Skeletal muscle;

RX MEDLINE=22388257; PubMed=12477932;

RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D., Altschul S.F., Zebberg B., Buetow K.H., Scheffer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hsieh F., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosack S.A., McSwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton B., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butcherfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schin J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [4]

RP SEQUENCE OF 54-408 FROM N.A.

RX MEDLINE=88199081; PubMed=2452201;

RA Sturgees A.D., Peterson M.G., McNeillage L.J., Whittingham S., Coppel R.S.;

RT "Characteristics and epitope mapping of a cloned human autoantigen La."

RL J. Immunol. 140:3212-3218 (1988).

RN [5]

RP SEQUENCE OF 54-97 FROM N.A.

RX MEDLINE=8516283; PubMed=3856888;

RA Chambers J.C., Keene J.D.;

RT "Isolation and analysis of cDNA clones expressing human lupus La antigen."

RL Proc. Natl. Acad. Sci. U.S.A. 82:2115-2119 (1985).

RN [6]

RP FUNCTION.

RX MEDLINE=89251617; PubMed=2470590;

RA Göttlieb E., Steltz J.A.;

RT "Function of the mammalian La protein: evidence for its action in transcription termination by RNA polymerase III."

RL EMBO J. 8:851-861 (1989).

RN [7]

RP PHOSPHORYLATION.

RX MEDLINE=97207017; PubMed=9054510;

RA Fan H., Sakulich A.L., Goodier J.L., Zhang X., Qin J., Marate R.J.;

RT "Phosphorylation of the human La antigen on serine 366 can regulate recycling of RNA polymerase III transcription complexes."

RL Cell 88:707-715 (1997).

RN [8]

RP INTERACTION WITH DDX15.

RX MEDLINE=22346609; PubMed=12458796;

RA Fouraux M.A., Kolkmann M.J.M., Van der Heijden A., De Jong A.S., Van Ventrcoijl W.J., Pruijn G.J.M.;

RT "The human La (SS-B) autoantigen interacts with DDX15/hPrp43, a putative DEAD-box RNA helicase."

RL RNA 8:1428-1443 (2002).

CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' termini of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including rRNA and 4.5S, 5S, 7S, and 7-2 RNAs.

CC -1- SUBUNIT: Interacts with DDX15.

CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).

CC -1- PTM: PHOSPHORYLATION. THE PHOSPHORYLATION SITES ARE AT THE C-TERMINAL PART OF THE PROTEIN.

CC -1- PTM: The N-terminus is blocked.

CC -1- DISEASE: Sera from patients with systemic lupus erythematosus often contain antibodies that react with the normal cellular La protein as if this antigen was foreign.

CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).

CC -----

DR EMBL: X13597; CAA31985.1; -.

DR EMBL: J04205; AA51885.1; -.

DR EMBL: BC001289; AA01289.1; -.

DR EMBL: BC020818; AA020818.1; -.

DR PIR: A11888; A11888.

DR Gene: HGNC:11316; SSB.

DR MIM: 109090;

DR GO: GO:0030529; C:ribonucleoprotein complex; TAS.

DR GO: GO:0003729; F:mRNA binding; TAS.

DR GO: GO:0000049; F:RNA binding; TAS.

DR GO: GO:0008334; P:histone mRNA metabolism; TAS.

DR GO: GO:0006400; P:RNA modification; TAS.

DR InterPro: IPR002344; Lupus_La.

DR InterPro: IPR006630; Lupus_La_dom.

DR InterPro: IPR000504; RNA_rec_mot.

DR Pfam: PF05383; La; 1.

DR Pfam: PF00076; rrm; 1.

DR PRINTS: PR00302; LUPUSLA.

DR SMART: SM00715; LA; 1.

DR SMART: SM00360; RRM; 1.

DR PROSITE: PS50102; RRM; 1.

DR PROSITE: PS00030; RRM_RNP_1; 1.

DR Systemic lupus erythematosus; RNA-binding; Phosphorylation; Nuclear protein.

FT DOMAIN 111 187 RNA-BINDING (RRM).

FT MOD_RES 366 366 PHOSPHORYLATION (BY CK2).

FT SEQUENCE 408 AA; 46837 MW; EC153C15F9187FC4 CRC64;

Query Match 46.3%; Score 44; DB 1; Length 408;

Best Local Similarity 50.0%; Pred. No. 11;

Mech. Dev. 37:69-80(1992).

RL MECH. Dev. 37:69-80(1992).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RA MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Ceinkner S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.H.C., Blazey R.G., Chame M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abrial J.F., Agbayani A., An H.-J., Andrews-Plankoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Bence P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,

RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahike C., Daveport L.B., Davies P.,

RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Gang N.S., Gelbart W.M., Glaeser K.,

RA Glodak A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai X.,

RA Lasco P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merklow G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K.A., Nusskern D.R., Pacleb J.M.,

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reiner K., Remington K., Sanders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of *Drosophila melanogaster*."

RL Science 287:2185-2195(2000).

CC -1- FUNCTION: May play a key role in egg organization. It may be a

CC transcriptional regulator.

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- TISSUE SPECIFICITY: Oocyte specific.

CC -1- DEVELOPMENTAL STAGE: Expressed at all oogenic stages.

CC -1- PTM: The N-terminus is blocked.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

CC EMBL, X63503; CAA45074.1; --

CC EMBL, AE003768; AAP56588.2; --

CC PIR, A56678; A56678.

DR FlyBase, FBgn0005536; yem-alpha.

DR Nuclear protein; DNA-binding; Repeat.

KW Nucleic acid binding; DNA-binding; Repeat.

FT DOMAIN 80 85 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

FT DOMAIN 207 217 POLY-SER.

FT DOMAIN 219 261 ASP/GUT-RICH (ACIDIC).

FT DOMAIN 230 253 2 X 12 AA TANDEM REPEATS.

FT REPEAT 230 241 1.

FT REPEAT 242 253 2.

FT REPEAT 242 253 2.

FT VARIANT 698 698 S -> L.

SO SEQUENCE 1002 AA; 109310 MW; EE69A384EBA24D2F CRC64;

Query Match 47.4%; Score 45; DB 1; Length 1002;

Best Local Similarity 52.9%; Pred. No. 19;

Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIIKOLEYFGN 17

DB 105 DDDVARIVLEAKYGN 121

RESULT 11

LA BOVIN

ID LA BOVIN STANDARD; PRT; 404 AA.

AC P10881;

DT 01-JUL-1989 (rel. 11, Created)

DT 01-JUL-1989 (rel. 11, Last sequence update)

DT 28-FEB-2003 (rel. 41, Last annotation update)

DE Lupa la protein homolog (La ribonucleoprotein) (La autoantigen

DE homolog).

GN SSB.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI_Taxid=9913;

RN [1]

RP SEQUENCE FROM N.A.

RP TISSUE=Plutitary;

RX MEDLINE=89202037; PubMed=2468131;

RA Chan E.K.L., Sullivan K.F., Tan E.M.;

RT "Ribonucleoprotein SSB/La belongs to a protein family with consensus

RT sequences for RNA-binding."

RL Nucleic Acids Res. 17:2233-2244(1989).

CC -1- FUNCTION: La protein plays a role in the transcription of RNA

CC polymerase III. It is most probably a transcription termination

CC factor. Binds to the 3' terminus of virtually all nascent

CC polymerase III transcripts. It is associated with precursor forms

CC of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S,

CC and 7-2 RNAs.

CC -1- SUBUNIT: Interacts with DDX15 (By similarity).

CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).

CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE

CC C-TERMINAL PART OF THE PROTEIN.

CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

CC EMBL, X13698; CAA11986.1; --

CC PIR, S03849; S03849.

DR InterPro, IPR002344; Lupa la.

DR InterPro, IPR006630; Lupa la dom.

DR InterPro, IPR000504; RNA_rec_mot.

DR Pfam, PF05383; La; 1.

DR Pfam, PF00076; rtm; 1.

DR PRINTS, PR00302; LUPUSLA.

DR SMART, SMO0715; LA; 1.

DR SMART, SMO0360; RRM; 1.

DR PROSITE, PS50102; RRM; 1.

DR PROSITE, PS00030; RRM_RNP_1; 1.

KW RNA-binding; Nuclear protein; Phosphorylation.

FT DOMAIN 111 187 RNA-BINDING (RRM).

SO SEQUENCE 404 AA; 46534 MW; 4EE30B5C262AD6A1 CRC64;

Query Match 46.3%; Score 44; DB 1; Length 404;

Best Local Similarity 50.0%; Pred. No. 11;

Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 DQRIKQLEYFGN 17


```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=2022256; PubMed=10761919;
RA Parthill J., Achman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagals K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491."
RL Nature 404:502-506(2000).
CC -1- CATALYTIC ACTIVITY: Diphosphate + H(2)O = 2 phosphate.
CC -1- COFACTOR: Binds 4 magnesium ions per subunit. Other metal ions can
CC support activity, but at a lower rate. Two magnesium ions are
CC required for the activation of the enzyme and are present before
CC substrate binds, two additional magnesium ions form complexes with
CC substrate and product (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the Pase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL16754; CAB64132.1; -.
DR PIR: G81930; G81930.
DR HSSP: P17288; IFMJ.
DR HAMAP: MF_00209; -.
DR InterPro: IPR008163; Inorg_pphsph.
DR InterPro: IPR008162; Pyrophosphatase.
DR Pfam: PF00719; Pyrophosphatase; 1.
DR ProDom: PD002014; Inorg_pphsph; 1.
DR PROSITE: PS00387; PPASE; 1.
DR HydroLase: Metal-binding; Magnesium; Complete proteome.
KW METAL. 67 67 MAGNESIUM 1 (BY SIMILARITY).
FT METAL. 72 72 MAGNESIUM 1 AND 2 (BY SIMILARITY).
FT METAL. 104 104 MAGNESIUM 1 (BY SIMILARITY).
SQ SEQUENCE 177 AA; 19799 MW; 1DAB8F80F8FC06F CRC64;

Query Match 47.4%; Score 45; DB 1; Length 177;
Best Local Similarity 43.8%; Pred. No. 3.3;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 DADQRIIKQLEYYFGN 17
DB 126 DLPOOLIKQIEHFHFNH 141

RESULT 9
ID IPYR_NEIMB STANDARD; PRT; 177 AA.
AC Q9K0G4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-
DE hydrolase) (Pase).
GN PPA OR NMB0641.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=2017575; PubMed=10710307;
RT Tetrelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,

```

```

RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Macon T., Clacko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
RA Cotton M.D., Uitterlivi T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Piza M., Grandi G., Sun L.,
RA Smith H.O., Frazer C.M., Moxon E.R., Rappunli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58."
RL Science 287:1809-1815(2000).
CC -1- CATALYTIC ACTIVITY: Diphosphate + H(2)O = 2 phosphate.
CC -1- COFACTOR: Binds 4 magnesium ions per subunit. Other metal ions can
CC support activity, but at a lower rate. Two magnesium ions are
CC required for the activation of the enzyme and are present before
CC substrate binds, two additional magnesium ions form complexes with
CC substrate and product (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the Pase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE002419; AAP41064.1; -.
DR PIR: F81175; F81175.
DR HSSP: P17288; IFMJ.
DR TIGR: NMB0641; -.
DR HAMAP: MF_00209; -.
DR InterPro: IPR008163; Inorg_pphsph.
DR InterPro: IPR008162; Pyrophosphatase.
DR Pfam: PF00719; Pyrophosphatase; 1.
DR ProDom: PD002014; Inorg_pphsph; 1.
DR PROSITE: PS00387; PPASE; 1.
DR HydroLase: Metal-binding; Magnesium; Complete proteome.
KW METAL. 67 67 MAGNESIUM 1 (BY SIMILARITY).
FT METAL. 72 72 MAGNESIUM 1 AND 2 (BY SIMILARITY).
FT METAL. 104 104 MAGNESIUM 1 (BY SIMILARITY).
SQ SEQUENCE 177 AA; 19811 MW; 6017182C446A8567 CRC64;

Query Match 47.4%; Score 45; DB 1; Length 177;
Best Local Similarity 43.8%; Pred. No. 3.3;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 DADQRIIKQLEYYFGN 17
DB 126 DLPOOLIKQIEHFHFNH 141

RESULT 10
ID YEMA_DROME STANDARD; PRT; 1002 AA.
AC P25992; Q9VAP4;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Yemanuclein-alpha.
GN YEM-ALPHA OR YEMA OR YG4.5 OR CG11879.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S;
RX MEDLINE=92297435; PubMed=1606021;
RX Ali-Ahmed O., Bellon B., Capri M., Joblet C., Thomas-Dejaage M.;
RT "The Yemanuclein-alpha; a new Drosophila DNA binding protein specific
for the oocyte nucleus."

```

```
CC      Name=2;
CC      IsoId=Q9Y5H7-2; Sequence=VSP_000679, VSP_000680;
CC
CC      Name=3;
CC      IsoId=Q9Y5H7-3; Sequence=VSP_008704, VSP_008705;
CC      Note=No experimental confirmation available;
CC      -1- SIMILARITY: Contains 6 cadherin domains.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
```

[illegible]

```

CC AC ID LA AEDAL STANDARD, PRT, 383 AA.
CC DT QZ64577
CC DT 15-JUL-1998 (Rel. 36, Created)
CC DT 15-JUL-1998 (Rel. 36, Last sequence update)
CC DE 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE La protein homolog (La ribonucleoprotein) (La autocantigen homolog) .
CC Aedes albopictus (Forest day mosquito) .
CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Prexygota;
CC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
CC OX NCBI_TaxID=7160;
CC RN [1]
CC RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
CC RX MEDLINE=96135233; PubMed=8551578;
CC RA Paragion N., Strauss U.H.;
CC RT "Mosquito homolog of the La autoantigen binds to Sindbis virus RNA.";
CC RL J. Virol. 70:1173-1181(1996) .
CC CC -I- FUNCTION: May be involved in transcription termination by RNA
      polymerase II. Binds RNA and DNA. Binds to the 3' end of the
      minus strand of Sindbis virus RNA. This may be significant for
      Sindbis virus RNA replication.
CC CC -I- SUBCELLULAR LOCATION: Nuclear. Primarily nuclear, but significant
      amounts are present in the cytoplasm.
CC CC -I- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC CC -I- SIMILARITY: TO VERTERBRATE PROTEIN LA.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as their content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).

```

```

DR  EMBL: S80954; AAB35931.1; -.
DR  InterPro: IPR002344; Lupus_La.
DR  InterPro: IPR006630; Lupus_La_dom.
DR  InterPro: IPR00504; RNA_rec_mot.
DR  Pfam: PF05383; La; 1.
DR  Pfam: PF00076; rrm; 1.
DR  PRINTS: PR00302; LUPUSLA.
DR  SMART: SM00715; La; 1.
DR  SMART: SM00360; RRM; 1.
DR  PROSITE: PS50102; RRM; 1.
DR  PROSITE: PS00030; RRM_RNP_1; FALSE NEG.
KW  RNA-binding; Nuclear protein; DNA-binding.
FT  DOMAIN 141 228 RNA-BINDING (RM).
SQ  SEQUENCE 383 AA; 44430 MW; 4B5CC8F21C40F452 CRC64;

Query Match
Best Local Similarity 49.5%; Score 47; DB 1; Length 383;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0
QY 8 IKOLEYFPGN 17
   |::|||||:
Db 48 IRQLEYFQSD 57

RESULT 8
IPYR_NEIMA
ID 1PYR_NEIMA STANDARD; PRT; 177 AA.
AC Q9JVG3;
DT 28-FEB-2003 (rel. 41, Created)
DT 28-FEB-2003 (rel. 41, Last sequence update)
DT 10-OCT-2003 (rel. 42, Last annotation update)
DE Inorganic pyrophosphatase (EC 3.6.1.1) (pyrophosphate phospho-
DE hydrolyase) (ppase).
DE PPA OR NMA0851.
GN Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCB1_TaxID=65699;

```

```

DR MIM: 601607; -
DR GO: GO:0005654; C:nucleoplasm; TAS.
DR GO: GO:0015074; P:DNA integration; TAS.
DR GO: GO:0006357; P:regulation of transcription from Pol II pro. .; TAS.
DR InterPro: IPR006939; SNFS.
DR Pfam: PF04855; SNFS; 1.
DR Transcription regulation; Activator; Nuclear protein;
DR Alternative splicing; Anti-oncogene.
FT VARSPLIC. 69 Missing (in isoform B).
FT CONFLICT 136 136 /FTid=VSP_004399.
FT CONFLICT 378 378 S -> P (IN REF. 2).
FT CONFLICT 382 382 L -> E (IN REF. 3).
FT CONFLICT 382 382 G -> A (IN REF. 2).
SQ SEQUENCE 385 AA; 44117 MW; B7BC6D6D5B443D CRC64;

Query Match 50.5%; Score 48; DB 1; Length 385;
Best Local Similarity 50.0%; Pred. No. 2.3;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Oy 1 DDADQRIIKQLEYFGNI 18
    :|||:|:|:|
Db 256 DOSDQRIIKINHVGNI 273

RESULT 5
SNFS_MOUSE STANDARD; PRT; 385 AA.
AC Q9Z0H3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SWI/SNF related, matrix associated, actin dependent regulator of
DE chromatin subfamily B member 1 (Integrase interactor 1 protein)
DE (msSNF5).
GN SMARCB1 OR SNF5L1 OR INI1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN NCB1
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
RX MEDLINE=99225325; PubMed=10208879;
RA Bruder C.E., Dumaniski J.P., Kedra D.;
RT "The mouse ortholog of the human SMARCB1 gene encodes two splice
RT forms";
RL Biochem. Biophys. Res. Commun. 257:886-890(1999).
CC -1- FUNCTION: Involved in chromatin-remodeling. Part of a complex that
CC opens the chromatin to facilitate the transcriptional machinery to
CC access their targets.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=A; Synonyms=INI1A;
CC Name=B; Synonyms=INI1B;
CC IsoId=Q9Z0H3-2; Sequence=VSP_004400;
CC -1- SIMILARITY: Belongs to the SNFS family.
CC -----
CC This SNFS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ011740; CA09761.1; -
DR EMBL: AJ011739; CA09760.1; -
DR MGD: MGI:132836; Smarcb1.
DR InterPro: IPR006939; SNFS.
DR Pfam: PF04855; SNFS; 1.
DR Transcription regulation; Activator; Nuclear protein;
KW Alternative splicing.

```

```

FT VARSPLIC 69 77 Missing (in isoform B).
FT SEQUENCE 385 AA; 44141 MW; B7BCA26875BD943D CRC64;
SQ SEQUENCE 385 AA; 44141 MW; B7BCA26875BD943D CRC64;

Query Match 50.5%; Score 48; DB 1; Length 385;
Best Local Similarity 50.0%; Pred. No. 2.3;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Oy 1 DDADQRIIKQLEYFGNI 18
    :|||:|:|:|
Db 256 DOSDQRIIKINHVGNI 273

RESULT 6
CDAS_HUMAN STANDARD; PRT; 936 AA.
ID CDAS_HUMAN
AC Q9Y5H7; Q75284; Q8NAR3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Protocadherin alpha 5 precursor (PCDH-alpha5).
DE PCDH5.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN NCB1
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=99306636; PubMed=10380929;
RA Wu Q., Mantatis T.;
RT "A striking organization of a large family of human neural cadherin-
RT like cell adhesion genes.";
RL Cell 97:779-790(1999).
RN NCB1
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Kimmerly W., Bondoc M., Cheng J., Connolly K.S., Gunning K.M.,
RA Kader K., Miguel T., Miller C., Pitluck S., Pollard M., Rojeski H.,
RA Subramanian S., Martin C.H.;
RT "Sequencing of human chromosome 5.";
RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
RN NCB1
RP SEQUENCE FROM N.A. (ISOFORM 3).
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Usdin T.B., Toshimiyuki S., Carrinot P., Prange C.,
RA Rosh S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bask S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield V.S.N., Krzywinski M.I., Skalek U., Smallus D.E.,
RA Schermer A., Schein J.E., Jones S.J.W., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Potential calcium-dependent cell-adhesion protein. May
CC be involved in the establishment and maintenance of specific
CC neuronal connections in the brain.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q9Y5H7-1; Sequence=Displayed;

```

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palczolo M., Pittman G.S., Pan S., Pollard J., Put V., Reese M.G.,
RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpston M., Skupski M.P., Smith T.,
RA Spier B., Spirdling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Massaman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of *Drosophila melanogaster*.";
RA Science 287:2185-2195(2000).
CC -1- FUNCTION: May be involved in transcription termination by RNA
CC polymerase III. Binds RNA and DNA. Binds to precursors of RNA
CC polymerase III transcripts. May play a specialized role during fly
CC development.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- DEVELOPMENTAL STAGE: Expressed throughout embryonic, larval,
CC pupal, and adult development. Expression throughout the embryo is
CC followed by a restricted pattern of mesodermal expression that is
CC later confined to the visceral mesoderm, gonads, gut, and salivary
CC glands.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -1- SIMILARITY: TO VERTERATE PROTEIN LA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U07652; AAA20518.1; -
CC EMBL; L32988; AAA21776.1; -
CC EMBL; AE003666; AAF5385.1; -
CC PIR; A53773; A53773.
CC PIR; A53781; A53781.
CC FlyBase; FBgn0011638; La.
CC GO; GO:0008098; F:5S rRNA primary transcript binding; IDA.
CC GO; GO:0003723; F:RNA binding; NAS.
CC InterPro; IPR002344; Lupus_La.
CC InterPro; IPR006630; Lupus_La_dom.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF00383; La; 1.
CC Pfam; PF00076; rrm; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00715; La; 1.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS50102; RRM; 1.
CC PROSITE; PS00030; RRM_RNP_1; 1.
CC KMW RNA-binding; Nuclear protein; DNA-binding.
CC FT DOMAIN 149 234 RNA-BINDING (RRM).
CC FT CONFLICT 169 169 A -> T (IN REF. 1).
CC FT CONFLICT 182 183 KH -> NS (IN REF. 1).
CC FT CONFLICT 283 283 A -> R (IN REF. 1).
CC FT CONFLICT 329 329 K -> N (IN REF. 1).
CC SEQUENCE 390 AA; 44884 MW; A8099288B90446A5 CRC64;

AC Q12824; O75784; O95474;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SWI/SNF related, matrix associated, actin dependent regulator of
DE chromatin subfamily B member 1 (Integrase interactor 1 protein)
DE (hsNF5) (BAF47).
DE SMARCB1 OR SNF5L1 OR INI1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95099327; PubMed=7801128;
RA Kalpana G.V., Marmon S., Wang W., Crabtree G.R., Goff S.P.;
RT "Binding and stimulation of HIV-1 integrase by a human homolog of
RT yeast transcription factor SNF5.";
RL Science 266:2002-2006 (1994).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=99225325; PubMed=10208879;
RA Bruder C.E., Dumanski J.P., Kedra D.;
RT "The mouse ortholog of the human SMARCB1 gene encodes two splice
RT forms.";
RL Biochem. Biophys. Res. Commun. 257:886-890 (1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98334382; PubMed=9671307;
RA Versteeg E.I., Sevenet N., Lange J., Rousseau-Merck M.F., Ambros P.,
RA Handgretinger R., Aurias A., Delattre O.;
RT "Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer.";
RL Nature 394:203-206 (1998).
CC -1- FUNCTION: Involved in chromatin-remodeling. Part of a complex that
CC opens the chromatin to facilitate the transcriptional machinery to
CC access their targets.
CC -1- SUBUNIT: Binds tightly to the human immunodeficiency virus-type 1
CC (HIV-1) integrase in vitro and stimulates its DNA-joining
CC activity.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=A; Synonyms=INI1A;
CC IsoId=Q12824-1; Sequence=displayed;
CC Name=B; Synonyms=INI1B;
CC IsoId=Q12824-2; Sequence=VSP_004399;
CC -1- DISEASE: TUMOR SUPPRESSOR. INACTIVATED IN MALIGNANT RHABDOID
CC TUMORS (MRT). MRT'S ARE EXTREMELY AGGRESSIVE CANCERS OF EARLY
CC CHILDHOOD.
CC -1- SIMILARITY: Belongs to the SNF5 family.
CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/SMARCB1ID169.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U04847; AAA81905.1; -
CC EMBL; AJ011738; AAA09759.1; -
CC EMBL; Y10118; CAA76639.1; ALU_SRO.
CC EMBL; Y17118; CAA76639.1; JOINED.
CC EMBL; Y17120; CAA76639.1; JOINED.
CC EMBL; Y17121; CAA76639.1; JOINED.
CC EMBL; Y17122; CAA76639.1; JOINED.
CC EMBL; Y17123; CAA76639.1; JOINED.
CC EMBL; Y17124; CAA76639.1; JOINED.
CC EMBL; Y17125; CAA76639.1; JOINED.
CC EMBL; Y17126; CAA76639.1; JOINED.
CC Genew; HGNC:11103; SMARCB1.

Query Match 52.6%; Score 50; DB 1; Length 390;
Best Local Similarity 57.1%; Pred. No. 1.1;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 4 DORIKOLEYFPGN 17
Db 51 ERAIRKQERYFGD 64

RESULT 4
SNF5_HUMAN
ID SNF5_HUMAN STANDARD; PRT; 385 AA.

DR PROSITE; PS00030; RRM_RNP_1; 1.
 KM RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 111 203 RNA-BINDING (RRM).
 FT DOMAIN 196 212 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 316 332 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 SQ SEQUENCE 428 AA; 48864 MW; AEB3A38B7D2E3EC3 CRC64;
 Query Match 60.0%; Score 57; DB 1; Length 428;
 Best Local Similarity 56.2%; Pred. No. 0.082;
 Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 QY 2 DADQRIIKOLEYFGN 17
 Db 12 DSDTKICEQIEYFGD 27

RESULT 2
 LAB_XENLA STANDARD; PRT; 427 AA.
 AC P28049;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus la protein homolog B (La ribonucleoprotein B) (La autoantigen homolog B).
 GN LAB1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8335;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocyte;
 RX MEDLINE=93287095; PubMed=8510143;
 RA Scherly D., Stutz F., Lin-Marq N., Clarkson S.G.;
 RT "La proteins from Xenopus laevis. cDNA cloning and developmental expression.";
 RT J. Mol. Biol. 231:196-204(1993).
 RU J. Mol. Biol. 231:196-204(1993).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' terminal of virtually all nascent polymerase III transcripts (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DEVELOPMENTAL STAGE: Barely detectable in stage I/II oocytes, accumulate in stage III/IV oocytes, then exhibit a roughly constant steady state level in mature oocytes, eggs, and early embryos.
 CC -1- PTM: Phosphorylated (Probable).
 CC -1- MISCELLANEOUS: There are two forms of La, Laa and Lab, in xenopus.
 CC -1- SIMILARITY: HIGH; TO MAMMALIAN LA PROTEIN.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X68818; CAA48716.1; -
 DR EMBL; X33817; S33817.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR00630; Lupus_La_dom.
 DR InterPro; IPR000504; RRM_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.

KM RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 110 202 RNA-BINDING (RRM).
 FT DOMAIN 315 331 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 SQ SEQUENCE 427 AA; 48995 MW; 45F3146F8934A355 CRC64;
 Query Match 57.9%; Score 55; DB 1; Length 427;
 Best Local Similarity 56.2%; Pred. No. 0.18;
 Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 2 DADQRIIKOLEYFGN 17
 Db 11 DLDTKICEQIEYFGD 26

RESULT 3
 LA_DROME STANDARD; PRT; 390 AA.
 AC P40796; Q24375; Q9VIN2;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 GN LA OR CG10922.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Canton-S; TISSUE=Ovary;
 RX MEDLINE=94309632; PubMed=8035794;
 RA Bai C., Li Z., Tolias P.P.;
 RT "Developmental characterization of a Drosophila RNA-binding protein homologous to the human systemic lupus erythematosus-associated La/SS-B autoantigen.";
 RT Mol. Cell. Biol. 14:5123-5129(1994).
 RU Mol. Cell. Biol. 14:5123-5129(1994).
 CC -1- SEQUENCE FROM N.A.
 CC MEDLINE=94309661; PubMed=8035818;
 CC Yoo C.-J., Wolin S.L.;
 CC "La proteins from Drosophila melanogaster and Saccharomyces cerevisiae: a yeast homolog of the La autoantigen is dispensable for growth.";
 CC Mol. Cell. Biol. 14:5412-5424(1994).
 CC [3]
 CC SEQUENCE FROM N.A.
 CC STRAIN=Berkely;
 CC MEDLINE=20196006; PubMed=10731132;
 CC Adams M.D., Celihner S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abril J.F., Aghayani A., An H.-J., Andrews-Ffannkoch C., Baldwin D., Balow R.M., Baer A., Baxendale J., Bayraktarglu L., Beasley E.M., Beeson K.T., Benos P.V., Berman B.P., Bhandari D., Boltsakov S., Bokova D., Botchan M.R., Bouck J., Brokstein P., Brottler P., Buttis J.M., Buam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferriz C., Fertler S., Fleischmann W., Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., Jatala M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.S., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z., Laake P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:01:41; Search time 7.24022 Seconds
(without alignments)
129.452 Million cell updates/sec

Title: US-09-836-073-17

Perfect score: 95

Sequence: 1 DDADRRITKQLEYFGNI 18

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	60.0	428	1	LAA_XENILA
2	55	57.9	427	1	P28048 xenopus lae
3	50	52.6	390	1	LAB_XENILA
4	48	50.5	385	1	LA_DROME
5	48	50.5	385	1	SNF5_HUMAN
6	48	50.5	385	1	SNF5_MOUSE
7	47	49.5	383	1	CDAS_HUMAN
8	45	47.4	177	1	IPYR_NEIMA
9	45	47.4	177	1	IPYR_NEIMA
10	45	47.4	1002	1	YEMA_DROME
11	44	46.3	404	1	LA_BOVIN
12	44	46.3	408	1	LA_HUMAN
13	44	46.3	415	1	LA_MOUSE
14	44	46.3	415	1	LA_MOUSE
15	44	46.3	2283	1	DPOE_MOUSE
16	43	45.3	176	1	IPYR_HAERIN
17	43	45.3	1290	1	IPYR_MYCPU
18	42	44.2	183	1	IPYR_MYCPU
19	42	44.2	214	1	EOTA_ACTHIO
20	42	44.2	365	1	KADC_MAIZE
21	42	44.2	222	1	ROAL_DROME
22	42	44.2	880	1	CADF_XENILA
23	42	44.2	884	1	CADF_XENILA
24	41.5	43.7	729	1	KEXI_YEAST
25	41	43.2	178	1	COAD_ZYMMO
26	41	43.2	211	1	TRMB_ZYMMO
27	41	43.2	214	1	KAD_MYCPU
28	41	43.2	298	1	LAHI_SCHPO
29	41	43.2	315	1	H02_RAT
30	41	43.2	350	1	Y535_SUISO
31	41	43.2	361	1	YPDF_ECOLI
32	41	43.2	466	1	SYH_BIFLO
33	41	43.2	1170	1	SMC2_YEAST

34	41	43.2	1292	1	RPOC_MYCCE
35	40	42.1	108	1	Y165_UREPA
36	40	42.1	339	1	DG11_THETN
37	40	42.1	500	1	SYA_RHILLO
38	40	42.1	594	1	SYA_BORBU
39	40	42.1	618	1	DNK_STRGR
40	40	42.1	832	1	TENL_CANAL
41	40	42.1	1132	1	NTU1_YEAST
42	39	41.1	106	1	RBS_CYAPA
43	39	41.1	139	1	RBS_GUITH
44	39	41.1	180	1	RBSA_SOLITU
45	39	41.1	180	1	RBSB_SOLITU

ALIGNMENTS

RESULT 1
ID LAA_XENILA STANDARD; PRT; 428 AA.
AC P28048;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Lupus La protein homolog A (La ribonucleoprotein A) (La autoantigen homolog A).
GN LAA1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RA MEDLINE=93287095; PubMed=8510143;
RX Scherly D., Stutz F., Lin-Marg N., Clarkson S.G.;
RT "La proteins from Xenopus laevis. cDNA cloning and developmental expression.";
RT J. Mol. Biol. 231:196-204(1993).
RL
CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' terminus of virtually all nascent polymerase III transcripts (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- DEVELOPMENTAL STAGE: Barely detectable in stage I/II oocytes, accumulate in stage III/IV oocytes, then exhibit a roughly constant steady state level in mature oocytes, eggs, and early embryos.
CC -1- PTM: Phosphorylated (Probable).
CC -1- MISCELLANEOUS: There are two forms of La, LAA and LAB, in xenopus.
CC -1- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL, X68817; CAA48715.1; -.
CC PIR, S33818; S33818.
CC InterPro: IPR002344; Lupus La.
CC InterPro: IPR006630; Lupus La dom.
CC InterPro: IPR000504; RNA_rec_mot.
CC Pfam: PF05383; La; 1.
CC Pfam: PF00076; Rrm; 1.
CC PRINTS: PR00302; LUPUSLA.
CC SMART: SM00715; LA; 1.
CC SMART: SM00360; RRM; 1.
CC PROSITE: PSS0102; RRM; 1.

A:Reference number: AB1775; MUID:20222556; PMID:10761919
A:Accession: G81930
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-177 <PAR>
A:Cross-references: GB:AL162754; GB:AL157959; NID:97379424; PIDN:CA884132.1; PID:9737956
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: ppa; NMA0851
C:Superfamily: inorganic pyrophosphatase
C:Keywords: hydrolase

Query Match 47.4%; Score 45; DB 2; Length 177;
Best Local Similarity 43.8%; Pred. No. 6.4;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 2 DADGRIIKOLEYFGN 17
Db 126 DLPOOLIKQIEHFHFNH 141

RESULT 12
A56678
Yemanuclein-alpha - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*
C:Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 21-Jun-2002
C:Accession: A56678; S22146
R:Alt-Alied: O.; Bellon, B.; Capri, M.; Uoblet, C.; Thomas-DeLaage, M.
Mech. Dev. 37, 69-80, 1992
A:Title: The yemanuclein-alpha: a new *Drosophila* DNA binding protein specific for the oc
A:Reference number: A56678; MUID:92297435; PMID:1606021
A:Accession: A56678
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1002 <AIE>
A:Cross-references: GB:X63503; NID:98837; PID:98838
C:Genetics:
A:Gene: FlyBase:yem-alpha
A:Cross-references: FlyBase:FBgn0005596
A:Introns: 80/3; 154/3; 428/1 477/2; 557/2
C:Keywords: DNA binding; oocyte

Query Match 47.4%; Score 45; DB 2; Length 1002;
Best Local Similarity 52.9%; Pred. No. 42;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DDADGRIIKOLEYFGN 17
Db 105 DDDVARIVKELEAKYGN 121

RESULT 13
F64526
adenine/cytosine DNA methyltransferase - *Helicobacter pylori* (strain 26695)
C:Species: *Helicobacter pylori*
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C:Accession: F64526
R:Tomd, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodok, A.; McKenna, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watney, L. Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.
A:Reference number: A64520; MUID:9739467; PMID:9252185
A:Accession: F64526
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-823 <TOM>
A:Cross-references: GB:AE000527; GB:AE000511; NID:92313131; PIDN:AA007124.1; PID:9231313

Query Match 46.8%; Score 44.5; DB 2; Length 823;
Best Local Similarity 42.1%; Pred. No. 41;
Matches 8; Conservative 6; Mismatches 2; Indels 3; Gaps 1;

Qy 2 DADGRIIKOLEYFGN 17
Db 65 ETKEIKOLEYFSSKFGN 83

RESULT 14
AD2410
hypothetical protein alr4836 [imported] - *Nostoc* sp. (strain PCC 7120)
C:Species: *Nostoc* sp. PCC 7120
A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AD2410
R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Itiguchi, N.; Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena*
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AD2410
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-158 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA076535.1; PID:917133973; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr4836

Query Match 46.3%; Score 44; DB 2; Length 158;
Best Local Similarity 50.0%; Pred. No. 8.3;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 5 ORIKOLEYFGN 18
Db 19 ORLIKSPQYWTGNL 32

RESULT 15
A13101
transcription regulator, *lyrR* family Atu4437 [imported] - *Agrobacterium tumefaciens* (str
C:Species: *Agrobacterium tumefaciens*
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C:Accession: A13101
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L. erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell. ; Karp, P.; Romero, P.; Zhang, S. Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, J. ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743393
A:Accession: A13101
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-300 <KUR>
A:Cross-references: GB:AE008689; PIDN:AL45231.1; PID:917742913; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu4437
A:Map position: linear chromosome
C:Superfamily: probable transcription regulator *lyrR*

Query Match 46.3%; Score 44; DB 2; Length 300;
Best Local Similarity 60.0%; Pred. No. 17;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 ADGRIIKOLEYFGN 17
Db 31 AVSRQIENLEYYFGS 45

Search completed: September 10, 2004, 18:02:37
Job time : 16.0782 secs

F:151-158/Region: RNA-binding RNP1 motif
F:228-428/Domain: phosphorylated #status predicted <PHY>

Query Match 60.0%; Score 57; DB 1; Length 428;
Best Local Similarity 56.2%; Pred. No. 0.18;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Oy 2 DADRIIKOLEYFGN 17
Db 12 DSDTKICEQIEYFGD 27

RESULT 3

S33817
ribonucleoprotein Ia-B - African clawed frog
N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
C:Species: Xenopus laevis (African clawed frog)
C>Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C:Accession: S33817; S28544
R:Schery, D.; Stutz, F.; Lin-Marg, N.; Clarkson, S.G.
J. Mol. Biol. 231, 196-204, 1993

A:Title: La proteins from *Xenopus laevis*. cDNA cloning and developmental expression.
A:Reference number: S33817; MUID:93287095; PMID:8510143
A:Accession: S33817

A:Molecule type: mRNA
A:Residues: 1-427 <SCH>

A:Cross-references: EMBL:G68618; NID:G64875; PIDN:CAA48716.1; PID:G64876
A:Comment: This protein associates with a variety of small RNA molecules, most of which
act as a transcription termination factor.

C:Superfamily: ribonucleoprotein Ia; ribonucleoprotein repeat homology

C:Keywords: phosphoprotein; RNA binding

F:111-177/Domain: ribonucleoprotein repeat homology <RNM>

F:112-117/Region: RNA-binding RNP2 motif

F:150-157/Region: RNA-binding RNP1 motif

F:227-427/Domain: phosphorylated #status predicted <PHY>

Query Match 57.9%; Score 55; DB 1; Length 427;
Best Local Similarity 56.2%; Pred. No. 0.37;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Oy 2 DADRIIKOLEYFGN 17
Db 11 DSDTKICEQIEYFGD 26

RESULT 4

T00677
hypothetical protein At2g43970 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein FEEL3.10
C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 23-Mar-2001
C:Accession: T00677; G84872
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,

submitted to the EMBL Data Library, June 1998

A:Description: Arabidopsis thaliana chromosome II BAC FEEL3 genomic sequence.

A:Reference number: Z14180

A:Accession: T00677

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-529 <ROU>

A:Cross-references: EMBL:AC004005; NID:G3212846; PID:G3212854

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.

ews, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J
Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: G84872

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-529 <STO>

A:Cross-references: GB:AB002093; NID:G3212854; PIDN:AAC23405.1; GSPDB:GN00139

C:Genetics:
A:Gene: FEEL3.10; At2g43970
A:Map:position: 2

A:introns: 200/3; 228/1; 259/3; 283/3; 310/2; 344/3; 367/3; 421/3

Query Match 56.8%; Score 54; DB 2; Length 529;
Best Local Similarity 44.4%; Pred. No. 0.63;
Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Oy 1 DDADRIIKOLEYFGNI 18
Db 191 EDSIOKIVNQYEVFSD 208

RESULT 5

A53773
La/SS-B homolog D-1a - fruit fly (*Drosophila melanogaster*)

C:Species: *Drosophila melanogaster*

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000

C:Accession: A53773

R:Bat, C.; Li, Z.; Tollas, P.P.

Mol. Cell. Biol. 14, 5123-5129, 1994

A:Title: Developmental characterization of a *Drosophila* RNA-binding protein homologous t

A:Reference number: A53773; MUID:94309632; PMID:8035794

A:Accession: A53773

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-390 <BAI>

A:Cross-references: GB:U07652; NID:G464019; PIDN:AAA20518.1; PID:G464020

C:Genetics:

A:Gene: FlyBase:la

A:Cross-references: FlyBase:FBgn0011638

C:Superfamily: ribonucleoprotein Ia; ribonucleoprotein repeat homology

C:Keywords: leucine zipper; RNA binding

Query Match 52.6%; Score 50; DB 2; Length 390;
Best Local Similarity 57.1%; Pred. No. 2.3;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Oy 4 DORIIKOLEYFGN 17
Db 51 ERATIRQYEVFGD 64

RESULT 6
A53781
ribonucleoprotein Ia - fruit fly (*Drosophila melanogaster*)

C:Species: *Drosophila melanogaster*

C>Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 20-Sep-1999

C:Accession: A53781

R:Yoo, C.J.; Molin, S.L.

Mol. Cell. Biol. 14, 5412-5424, 1994

A:Title: La proteins from *Drosophila melanogaster* and *Saccharomyces cerevisiae*: a yeast t

A:Reference number: A53781; MUID:94309661; PMID:8035818

A:Accession: A53781

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-390 <YOO>

A:Cross-references: GB:L32988; NID:G488469; PID:G488470

C:Genetics:

A:Gene: FlyBase:la

A:Cross-references: FlyBase:FBgn0011638

C:Superfamily: ribonucleoprotein Ia; ribonucleoprotein repeat homology

C:Keywords: RNA binding

Query Match 52.6%; Score 50; DB 2; Length 390;
Best Local Similarity 57.1%; Pred. No. 2.3;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Oy 4 DORIIKOLEYFGN 17
Db 51 ERATIRQYEVFGD 64

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:33:11; Search time 14.0782 Seconds
(without alignments)
122.988 Million cell updates/sec

Title: US-09-836-073-17

Perfect score: 95

Sequence: 1 DDADQRIIKOLEYVFGNI 18

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	92	96.8	396	2	T30953	hypothetical prote
2	57	60.0	428	1	S33818	ribonucleoprotein
3	55	57.9	427	1	S33817	ribonucleoprotein
4	54	56.8	529	2	T00677	hypothetical prote
5	50	52.6	390	2	A53773	La/SS-B homolog D-
6	50	52.6	390	2	A53781	ribonucleoprotein
7	48	50.5	385	2	S54705	transcription fact
8	46	48.4	391	2	H89777	capsular polysacchr
9	45	47.4	156	2	D75605	hypothetical prote
10	45	47.4	177	2	F81175	inorganic pyrophos
11	45	47.4	177	2	G81930	probable inorganic
12	45	47.4	1002	2	A56578	yemamuclein-alpha
13	44.5	46.8	823	2	F64526	adenine/cytosine D
14	44	46.3	158	2	AD2410	hypothetical prote
15	44	46.3	300	2	A13101	transcription regu
16	44	46.3	332	2	A96185	probable transcrip
17	44	46.3	334	2	G97972	conserved hypochet
18	44	46.3	404	1	S03849	ribonucleoprotein
19	44	46.3	408	1	A31888	ribonucleoprotein
20	44	46.3	415	1	JCI1494	ribonucleoprotein
21	44	46.3	455	2	G95104	hypothetical prote
22	43	45.3	176	2	G64049	inorganic pyrophos
23	43	45.3	342	2	A70183	hypothetical prote
24	43	45.3	780	2	T27659	hypothetical prote
25	43	45.3	1290	2	S73653	DNA-directed RNA p
26	42	44.2	183	2	F90570	hypothetical prote
27	42	44.2	222	2	S45634	adenylate kinase (
28	42	44.2	340	2	A89798	hypothetical prote
29	42	44.2	365	2	A26459	helix-destabilizin

30	42	44.2	840	2 B87467	conserved hypothet
31	42	44.2	895	1 IXLCP	EP-cadherin precur
32	42	44.2	905	2 S43064	cadherin - African
33	41.5	43.7	188	2 B97104	phospholipase D fa
34	41.5	43.7	729	2 A29651	KEX1 protein precu
35	41	43.2	178	2 S77850	probable histidine
36	41	43.2	192	1 S75273	hypothetical prote
37	41	43.2	214	2 G90582	adenylate kinase (
38	41	43.2	298	2 T38937	RNA-binding protei
39	41	43.2	298	2 T43542	RNA-binding protei
40	41	43.2	315	1 A35199	heme oxygenase (de
41	41	43.2	346	2 S78647	hypothetical prote
42	41	43.2	350	2 G90199	acyl carrier prote
43	41	43.2	361	1 F65012	hypothetical prote
44	41	43.2	434	2 AC0512	putative sulfatas
45	41	43.2	450	2 T21931	hypothetical prote

ALIGNMENTS

```

RESULT 1
T30953
hypothetical protein C44E4.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000
C:Accession: T30953
R:Sammons, L., Wohldmann, P.; Gillam, B.
Submitted to: The EMBL Data Library, August 1999
A:Description: The sequence of C. elegans com1d C44E4.
A:Reference number: Z20945
A:Accession: T30953
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-396 <SAM>
A:Cross-References: EMBL:AF003140; PIDN:AMB54169.1
A:Experimental source: strain Bristol N2; clone C44E4
C:Genetics:
A:Map position: 1
A:Note: C44E4.4
A:Introns: 45/1; 114/3
C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

Query Match          96.8%; Score 92; DB 2; Length 396;
Best Local Similarity 94.4%; Pred. No. 2.8e-07;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 DDADQRIIKOLEYVFGNI 18
      |||||:|||||
Db      10 DDADQRIIKOLEYVFGNI 27

RESULT 2
S33818
ribonucleoprotein La.A - African clawed frog
N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
C:Species: Xenopus laevis (African clawed frog)
C>Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C:Accession: S33818; S28545
R:Scherly, D.; Stutz, F.; Lin-Marg, N.; Clarkon, S.G.
J. Mol. Biol. 231, 196-204, 1993
A>Title: La proteins from Xenopus laevis. cDNA cloning and developmental expression.
A:Reference number: S33817; MID:93287095; PMID:8510143
A:Accession: S33818
A:Molecule type: mRNA
A:Residues: 1-428 <SCH>
A:Cross-References: EMBL:X68817; NID:G64873; PIDN:CAA48715.1; PID:G64874
C:Comment: This protein associates with a variety of small RNA molecules, most of which
ay act as a transcription termination factor.
C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
C:Keywords: phosphoprotein; RNA binding
F:112-178/Domain: ribonucleoprotein repeat homology <RRM>
F:113-118/Region: RNA-binding RNP2 motif

```

This Page Blank (usp10)

PA (ADHE-) ADHEREX TECHNOLOGIES INC.

XX Blaschuk OM, Gour BJ, Byers S;

DR WPI, 2000-038791/03.

XX
PT New cadherin modulating agents, used for modulating nonclassical cadherin-mediated functions for treating e.g. cancers, obesity, rheumatoid arthritis, multiple sclerosis, diabetes or a neurological disease.

XX
PS Disclosure; Fig 2; 252pp; English.

XX
CC The present invention describes cadherin modulating agents (MA) comprising peptides which comprise a nonclassical cadherin cell adhesion recognition (CAR) sequence. The MA can be used for modulating nonclassical cadherin-mediated functions. They can be used for e.g. inhibiting adhesion of nonclassical-cadherin expressing cells in a mammal, enhancing delivery of a drug through the skin of a mammal, enhancing delivery of a drug to a tumour in a mammal, treating cancer in a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-expressing cell, preventing or treating obesity in a mammal, stimulating blood vessel regression in a mammal, enhancing drug delivery to the central nervous system, treating a demyelinating neurological disease, increasing vasopermeability in a mammal, enhancing adhesion of nonclassical cadherin-expressing cells, inhibiting synaptic stability in a mammal, or preventing pregnancy in a mammal. They can also be used for e.g. enhancing or directing neurite outgrowth, facilitating wound healing or reducing scar tissue, or enhancing adhesion of foreign tissue in a mammal. They can also be used for treating e.g. psoriasis, arthritis, age-related macular degeneration, multiple sclerosis and diabetes. The products can also be used for detection and diagnosis and in bioreactors. AA60592 to AA64572 represent specifically claimed peptides, and CC AA64573 to AA64643 and AA63183 to AA63186 represent sequences used in the exemplification of the present invention

XX
SQ Sequence 108 AA:

Query Match 49.5%; Score 47; DB 3; Length 108;

Best Local Similarity 47.1%; Pred. No. 5.5;

Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 DADQRIKQLEYYEGNI 18

DB 26 DADEGINKETLTPFSNL 42

Search completed: September 10, 2004, 17:51:53
Job time : 59.4128 secs

XX 16-DEC-1999; 99UP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 PA (KYOWA) KYOWA HAKKO KOGYO KK.
 XX
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX
 DR MPI; 2001-376931/40.
 DR N-PSDB; AAH66529.
 XX
 PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 PS
 XX Claim 17; SEQ ID NO 5064; 246bp + Sequence Listing; English.
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Corynebacterium bacteria Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of corynebacterium bacteria, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Corynebacterium bacteria, and identifying a homologue of a gene derived from
 CC Corynebacterium bacteria. Corynebacterium bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described in the
 CC exemplification of the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from the European Patent Office
 XX
 SQ Sequence 397 AA;
 Query Match 50.5%; Score 48; DB 4; Length 397;
 Best Local Similarity 50.0%; Pred. NO. 17;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 2 DADQRIIKOLEYFGN 17
 Db 138 DAQCSILDRMEYFGN 153
 RESULT 14
 ABG01240
 ID ABG01240 standard; protein; 650 AA.
 XX
 AC ABG01240;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #1231.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 PI
 DR MPI; 2001-639362/73.
 DR N-PSDB; AAS65427.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 31599; 103bp; English.
 CC
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (II) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 650 AA;
 Query Match 50.5%; Score 48; DB 4; Length 650;
 Best Local Similarity 47.1%; Pred. NO. 30;
 Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 2 DADQRIIKOLEYFGN 18
 Db 223 DADGINKELIVFPSNL 239
 RESULT 15
 AAY64613
 ID AAY64613 standard; peptide; 108 AA.
 XX
 AC AAY64613;
 XX
 DT 02-MAR-2000 (first entry)
 XX
 DE Nonclassical cadherin extracellular domain SEQ ID NO:41.
 XX
 KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
 KW cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
 KW cadherin related neuronal receptor; LI-cadherin; procadherin;
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; osteoly;
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
 KW neurological disease.
 XX
 OS Mammalia.
 XX
 PN WO9957149-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 05-MAY-1999; 99WO-CA000363.
 XX
 PR 05-MAY-1998; 98US-00073040.
 PR 06-NOV-1998; 98US-00187859.
 PR 20-JAN-1999; 99US-00234395.
 PR 08-MAR-1999; 99US-00264516.
 XX

PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159633P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 56.8%; Score 54; DB 3; Length 993;
Best Local Similarity 44.4%; Pred. No. 4.8; Mismatches 0;
Matches 8; Conservative 7; Indels 3; Gaps 0;

OY 1 DDADQRIKOLEYYPFGN 18
DB 191 EDSIQKIVNQEYVFSDL 208

RESULT 7
ID ABG72119 standard; peptide; 16 AA.
AC ABG72119;
DT 28-JAN-2003 (first entry)
DE Viral replication inhibiting peptide, DROSOPHILA.
XX
XX Viral replication inhibitor; IRBS initiated translation; LAP;
KW internal ribosome entry site initiated translation; La antigen protein;
KM viral life cycle; antiviral therapy; pharmaceutical; veterinary;
XX agricultural; horticultural; vitucide.
OS
XX Drosophila sp.
XX
XX WO200283858-A2.
XX
PD 24-OCT-2002.
XX
PF 12-APR-2002; 2002WO-US011589.
XX
PR 16-APR-2001; 2001US-00836073.
XX
PA (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.
XX
PI Dasgupta A, Das S, Baidya N;
XX
XX WPI; 2003-058634/05.
XX
XX
PT New compound containing acidic and aromatic amino acids, useful as
PT antiviral therapy in pharmaceutical, veterinary or
PT agricultural/horticultural applications.
XX
XX Claim 1; Page 15; 19pp; English.
XX
XX The present invention relates to peptides and methods of inhibiting the
CC replication of viruses that utilise internal ribosome entry site (IRES)
CC initiated translation, and/or inhibiting viruses that utilise the La

CC antigen protein (LAP) in any phase of their life cycle. The peptides of
CC the invention compete with LAP and inhibit the utilisation of various
CC biochemical and physiological functions of LAP required for a productive
CC life cycle. The methods and compositions are useful as antiviral therapy
CC in pharmaceutical, veterinary or agricultural/horticultural applications.
CC ABG72101-ABG72119 represent peptides useful as antiviral agents
XX
SQ Sequence 16 AA;

Query Match 52.6%; Score 50; DB 6; Length 16;
Best Local Similarity 57.1%; Pred. No. 0.18; Mismatches 1; Indels 0; Gaps 0;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 4 DQRIKOLEYYPFGN 17
DB 2 ERAIRKQVEYFQSD 15

RESULT 8
ID ABB65316 standard; protein; 390 AA.
AC ABB65316;
XX
XX
XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 22740.

XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmacological.
XX
XX Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

PA (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL09419.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.

XX Disclosure; SEQ ID NO 22740; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 390 AA;

Query Match 52.6%; Score 50; DB 4; Length 390;

Best Local Similarity 57.1%; Pred. No. 7.7; Mismatches 1; Indels 0; Gaps 0;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 4 DQRIKOLEYYPFGN 17

PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130691P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 01-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 15-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147955P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149733P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0150656P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.

PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144336P.
PR 20-JUL-1999; 99US-0144337P.
PR 20-JUL-1999; 99US-0144338P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 22-SEP-1999; 99US-0155466P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.

PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 56.8%; Score 54; DB 3; Length 923;

Best Local Similarity 44.4%; Pred. No. 4.4; Mismatches 3; Indels 0; Gaps 0;

Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 DDADRIIKOLEYFQNI 18
DB 121 EDSDIKVQVEYFEDL 138

RESULT 6
ID AAG47712 standard; protein; 993 AA.
AC AAG47712;
XX 18-OCT-2000 (First entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 60166.
DE Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EPI033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.

PR 24-SEP-1999; 99US-0156569P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159633P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 56.8%; Score 54; DB 3; Length 913;
Best Local Similarity 44.4%; Pred. No. 4.4;
Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Cy i DDADRRITKOLEYFGNI 18
Db 111 EDSITQIVNQVEYFSDL 128

RESULT 5
AAG47713

ID AAG47713 standard; protein; 923 AA.

AC AAG47713;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO. 60167.

KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

OS Arabidopsis thaliana.

PN BP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 21-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136788P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.

XX 25-FEB-2000; 2000EP-00301439.
PF
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 04-MAY-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0134941P.
PR 21-MAY-1999; 99US-0135124P.
PR 24-MAY-1999; 99US-0135353P.
PR 25-MAY-1999; 99US-0135629P.
PR 27-MAY-1999; 99US-0136021P.
PR 28-MAY-1999; 99US-0136392P.
PR 01-JUN-1999; 99US-0136782P.
PR 03-JUN-1999; 99US-0137222P.
PR 04-JUN-1999; 99US-0137528P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.

PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147416P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147933P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149168P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0150566P.
PR 25-AUG-1999; 99US-0150884P.
PR 26-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151388P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153707P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.

Query Match 100.0%; Score 95; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 5e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DDADRIIKOLEYFGN 18
 |||||
 DB 1 DDADRIIKOLEYFGN 18

RESULT 2

ABP70526
 ID ABP70526 standard; protein; 411 AA.

AC ABP70526;

DT 22-APR-2003 (first entry)

DE Histone deacetylase chromatin-associated protein complex subunit MYC.

KW Histone acetyltransferase; histone deacetylase; gene expression profile;

KM chromatin-associated protein; gene expression.

XX Unidentified.

OS WO2003000715-A1.

PN 03-JAN-2003.

XX 21-JUN-2002; 2002MO-US019750.

PR 22-JUN-2001; 2001US-0300135P.

XX (CERE-) CERES INC.

PI Dang V, Okamoto J;

DR WPI; 2003-175280/17.

XX New chimeric polypeptide comprising a histone acetyltransferase

PT polypeptide segment and a segment comprising a histone deacetylase

PT chromatin-associated protein complex subunit, useful for modulating gene

PT expression in cells.

XX Disclosure; Page 61; 85pp; English.

XX The specification describes chimeric histone acetyltransferase

CC polypeptides. The chimeric polypeptides comprise a polypeptide segment

CC that exhibits histone acetyltransferase activity, and a polypeptide

CC segment having 40% or greater sequence identity to a subunit of a histone

CC deacetylase chromatin-associated protein complex. The chimeric

CC polypeptides are useful for determining gene expression profiles in

CC specific cells, for modulating gene expression in specific cells, and for

CC making genetically modified eukaryotes. The present sequence represents

CC an exemplary histone deacetylase chromatin-associated protein complex

CC subunit, which can be used to produce chimeric polypeptides of the

CC invention

RESULT 3

ABG72115
 ID ABG72115 standard; peptide; 18 AA.

AC ABG72115;

DT 28-JAN-2003 (first entry)

DE Viral replication inhibiting peptide, XENOPUS.

KW Viral replication inhibitor; IRES initiated translation; IAP;

KM internal ribosome entry site initiated translation; Ia antigen protein;

KW viral life cycle; antiviral therapy; pharmaceutical; veterinary;

XX agricultural; horticultural; virucide; xenopus.

OS Xenopus sp.

PN WO200283858-A2.

XX 24-OCT-2002.

PF 12-APR-2002; 2002MO-US011589.

PR 16-APR-2001; 2001US-00836073.

XX (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.

PI Dasgupta A, Das S, Balda N;

DR WPI; 2003-058634/05.

XX New compound containing acidic and aromatic amino acids, useful as

PT antiviral therapy in pharmaceutical, veterinary or

PT agricultural/horticultural applications.

XX Claim 1; Page 15; 19pp; English.

XX The present invention relates to peptides and methods of inhibiting the

CC replication of viruses that utilize internal ribosome entry site (IRES)

CC initiated translation, and/or inhibiting viruses that utilize the Ia

CC antigen protein (IAP) in any phase of their life cycle. The peptides of

CC the invention compete with IAP and inhibit the utilisation of various

CC biochemical and physiological functions of IAP required for a productive

CC life cycle. The methods and compositions are useful as antiviral therapy

CC in pharmaceutical, veterinary or agricultural/horticultural applications.

CC ABG72101-ABG72119 represent peptides useful as antiviral agents

CC Sequence 18 AA;

XX Query Match 57.9%; Score 55; DB 6; Length 18;

XX Best Local Similarity 56.2%; Pred. No. 0.03;

XX Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 2 DDADRIIKOLEYFGN 17
 |||||

DB 2 DLDTRICOLEYFGD 17

RESULT 4

AAG47714
 ID AAG47714 standard; protein; 913 AA.

AC AAG47714;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 60168.

KW Protein identification; signal transduction pathway; metabolic pathway;

KM hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 16:58:20 ; Search time 56.3129 Seconds
(without alignments)
90.314 Million cell updates/sec

Title: US-09-836-073-17

Perfect score: 95
Sequence: 1 DDADQRIIKQLEYFGNI 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95	100.0	18	6	ABG72117
2	57	60.0	411	6	ABP70526
3	55	57.9	18	6	ABG72115
4	54	56.8	913	3	AA647714
5	54	56.8	923	3	AA647713
6	54	56.8	993	3	AA647712
7	50	52.6	16	6	ABG72119
8	50	52.6	390	4	ABG65316
9	48	50.5	376	4	ABG93709
10	48	50.5	384	3	AA923337
11	48	50.5	385	2	AAW00081
12	48	50.5	394	4	AA895622
13	48	50.5	397	4	AA893310
14	48	50.5	650	4	ABG01240
15	47	49.5	108	3	AA646613
16	47	49.5	176	6	AAE34638
17	46	48.4	381	6	ABJ19009
18	46	48.4	391	6	ABM72091
19	45	47.4	18	6	ABG72102
20	45	47.4	193	6	ABP77967
21	45	47.4	198	4	AAW25632
22	45	47.4	223	5	ABB80984
23	45	47.4	393	6	ABD10476
24	45	47.4	445	6	ABD10474
25	45	47.4	1002	4	ABB58137

26	45	47.4	1062	4	AAU01014	AAU01014	Human sec
27	45	47.4	1062	4	AA662238	AA662238	Human cal
28	45	47.4	1076	5	ABG95850	ABG95850	Human cal
29	45	47.4	1076	6	ABO19418	ABO19418	Human cal
30	45	47.4	1082	4	AAU01015	AAU01015	Human sec
31	45	47.4	1082	4	AA662239	AA662239	Human cal
32	45	47.4	1109	4	AAU01016	AAU01016	Human sec
33	45	47.4	1109	4	AA662240	AA662240	Human cal
34	45	47.4	1115	4	AAU01025	AAU01025	Human cal
35	45	47.4	1115	4	AA662249	AA662249	Human cal
36	45	47.4	1145	3	AA923322	AA923322	Human alp
37	45	47.4	1145	4	AAU01023	AAU01023	Human sec
38	45	47.4	1145	4	AA662247	AA662247	Human cal
39	45	47.4	1145	5	ABG95848	ABG95848	Human cal
40	45	47.4	1145	5	ABG95849	ABG95849	Human cal
41	45	47.4	1145	6	ABO19417	ABO19417	Human cal
42	45	47.4	1145	6	ABO19416	ABO19416	Human cal
43	44.5	46.8	823	2	AAW98507	AAW98507	H. pylori
44	44	46.3	17	6	ABG72113	ABG72113	Viral rep
45	44	46.3	18	3	AAV52200	AAV52200	Human la

ALIGNMENTS

RESULT 1	ABG72117	standard; peptide; 18 AA.
ID	ABG72117	standard; peptide; 18 AA.
AC	ABG72117	
DT	28-JAN-2003	(first entry)
DE	Viral replication inhibiting peptide, C. ELEGANS.	
XX	Viral replication inhibitor; IRES initiated translation; LAP;	
XX	internal ribosome entry site initiated translation; La antigen protein;	
XX	viral life cycle; antiviral therapy; pharmaceutical; veterinary;	
XX	agricultural; horticultural; virucide.	
OS	Caenorhabditis elegans.	
PN	W0200283858-A2.	
PD	24-OCT-2002.	
XX	12-APR-2002; 2002WO-US011589.	
PF	16-APR-2001; 2001US-00836073.	
PR	(UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.	
PA	Dasgupta A, Das S, Baidya N;	
XX	WPI; 2003-058634/05.	
PT	New compound containing acidic and aromatic amino acids, useful as	
PT	antiviral therapy in pharmaceutical, veterinary or	
PT	agricultural/horticultural applications.	
XX	Claim 1; Page 15; 19pp; English.	
PS	The present invention relates to peptides and methods of inhibiting the	
CC	replication of viruses that utilise internal ribosome entry site (IRES)	
CC	initiated translation, and/or inhibiting viruses that utilise the La	
CC	antigen protein (LAP) in any phase of their life cycle. The peptides of	
CC	the invention compete with LAP and inhibit the utilisation of various	
CC	biochemical and physiological functions of LAP required for a productive	
CC	life cycle. The methods and compositions are useful as antiviral therapy	
CC	in pharmaceutical, veterinary or agricultural/horticultural applications.	
CC	ABG72101-ABG72119 represent peptides useful as antiviral agents	
XX	Sequence 18 AA;	
SQ		

This Page Blank (uspto)

Mon Sep 13 09:36:14 2004

us-09-836-073-16.rapb

Page 5

Search completed: September 10, 2004, 18:11:56
Job time : 42.6642 secs

```
FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match      77.4%; Score 80.5; DB 9; Length 18;
Best Local Similarity 89.5%; Pred. No. 1e-05;
Matches 17; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Cy 1 AALEAKICHQIEEYFGDF 19
   ||||| ||||| |||||
Db 1 AALEAKICHQI-EYFGDF 18

RESULT 12
US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-10

Query Match      76.4%; Score 79.5; DB 9; Length 18;
Best Local Similarity 89.5%; Pred. No. 1.5e-05;
Matches 17; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Cy 1 AALEAKICHQIEEYFGDF 19
   ||||| ||||| |||||
Db 1 AALEAKICHQI-EYFGDF 18

RESULT 13
US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
```

```
US-09-836-073-2

Query Match      73.6%; Score 76.5; DB 9; Length 18;
Best Local Similarity 84.2%; Pred. No. 4.4e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Cy 1 AALEAKICHQIEEYFGDF 19
   ||||| ||||| |||||
Db 1 AALEAKICHQI-EYFGDF 18

RESULT 14
US-09-836-073-3
; Sequence 3, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-3

Query Match      73.6%; Score 76.5; DB 9; Length 18;
Best Local Similarity 78.9%; Pred. No. 4.4e-05;
Matches 15; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Cy 1 AALEAKICHQIEEYFGDF 19
   ||||| ||||| |||||
Db 1 AALQAKICHQI-QYFGQF 18

RESULT 15
US-09-836-073-4
; Sequence 4, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-4

Query Match      73.6%; Score 76.5; DB 9; Length 18;
Best Local Similarity 93.8%; Pred. No. 4.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Cy 4 EAKICHQIEEYFGDF 19
   ||||| ||||| |||||
Db 4 EAKICHQI-EYFGDF 18
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-806-695

Query Match
Best Local Similarity 85.1%; Score 88.5; DB 14; Length 460;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
Db 63 AALEAKICHQI-EYYFGDF 80

RESULT 7
US-10-264-049-2643
; Sequence 2643, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P4133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2643
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-2643

Query Match
Best Local Similarity 85.1%; Score 88.5; DB 15; Length 460;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
Db 63 AALEAKICHQI-EYYFGDF 80

RESULT 8
US-09-836-073-13
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baldiva, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Mouse
US-09-836-073-13

Query Match
Best Local Similarity 81.2%; Score 84.5; DB 9; Length 17;
Matches 17; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2 AALEAKICHQIEYYFGDF 19
Db 1 AALEAKICHQI-EYYFGDF 17
```

```

RESULT 9
US-09-836-073-11
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baldiva, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-073-11

Query Match
Best Local Similarity 80.3%; Score 83.5; DB 9; Length 18;
Matches 17; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
Db 1 AALEAKICHQIEQ-YFGDF 18

RESULT 10
US-09-836-073-9
; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baldiva, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-073-9

Query Match
Best Local Similarity 79.3%; Score 82.5; DB 9; Length 18;
Matches 17; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGD 18
Db 1 AALEAKICHQI-EYYFGD 17

RESULT 11
US-09-836-073-12
; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baldiva, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
```

```

; APPLICANT: Balda, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-1

Query Match      85.1%; Score 88.5; DB 9; Length 18;
Best Local Similarity 94.7%; Pred. No. 5.7e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY      1 AALEAKICHQIEYYFGDF 19
      |||||||
DB      1 AALEAKICHQI-EYYFGDF 18

RESULT 3
US-09-836-073-14
; Sequence 14, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balda, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bovine
US-09-836-073-14

Query Match      85.1%; Score 88.5; DB 9; Length 18;
Best Local Similarity 94.7%; Pred. No. 5.7e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY      1 AALEAKICHQIEYYFGDF 19
      |||||||
DB      1 AALEAKICHQI-EYYFGDF 18

RESULT 4
US-10-170-385-477
; Sequence 477, Application US/10170385
; Publication No. US20030203372A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 532682000100
; CURRENT APPLICATION NUMBER: US/10/170,385
```

```

; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 477
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-170-385-477

Query Match      85.1%; Score 88.5; DB 12; Length 408;
Best Local Similarity 94.7%; Pred. No. 1.5e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY      1 AALEAKICHQIEYYFGDF 19
      |||||||
DB      11 AALEAKICHQI-EYYFGDF 28

RESULT 5
US-09-925-298-695
; Sequence 695, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 695
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-298-695

Query Match      85.1%; Score 88.5; DB 12; Length 460;
Best Local Similarity 94.7%; Pred. No. 1.7e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY      1 AALEAKICHQIEYYFGDF 19
      |||||||
DB      63 AALEAKICHQI-EYYFGDF 80

RESULT 6
US-10-102-806-695
; Sequence 695, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 695
; LENGTH: 460
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:52:06 ; Search time 42.5642 Seconds
(without alignments)
143.151 Million cell updates/sec

Title: US-09-836-073-16

Perfect score: 104

Sequence: 1 AALEAKICHQIEHYFGDF 19

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	100.0	19	9	US-09-836-073-16
2	88.5	85.1	18	9	US-09-836-073-1
3	88.5	85.1	18	9	US-09-836-073-14
4	88.5	85.1	408	12	US-10-170-385-477
5	88.5	85.1	460	12	US-09-925-298-695
6	88.5	85.1	460	14	US-10-102-806-695
7	88.5	85.1	460	15	US-10-264-049-2643
8	84.5	81.2	17	9	US-09-836-073-13
9	83.5	80.3	18	9	US-09-836-073-11
10	82.5	79.3	18	9	US-09-836-073-9
11	80.5	77.4	18	9	US-09-836-073-12
12	79.5	76.4	18	9	US-09-836-073-10
13	76.5	73.6	18	9	US-09-836-073-2
14	76.5	73.6	18	9	US-09-836-073-3
15	76.5	73.6	18	9	US-09-836-073-4

16	75.5	72.6	18	9	US-09-836-073-8	Sequence 8, Appli
17	73.5	70.7	18	9	US-09-836-073-7	Sequence 7, Appli
18	65.5	63.0	18	9	US-09-836-073-15	Sequence 15, Appli
19	64.5	62.0	18	9	US-09-836-073-5	Sequence 5, Appli
20	63	60.6	18	9	US-09-836-073-6	Sequence 6, Appli
21	62.5	60.1	21	15	US-10-376-121A-20	Sequence 20, Appli
22	58	55.8	37	9	US-09-843-676-24	Sequence 24, Appli
23	58	55.8	37	9	US-09-766-253-24	Sequence 24, Appli
24	58	55.8	37	10	US-09-438-486-24	Sequence 24, Appli
25	58	55.8	37	14	US-10-053-758-25	Sequence 24, Appli
26	58	55.8	37	14	US-10-054-225-24	Sequence 24, Appli
27	58	55.8	37	14	US-10-054-611-24	Sequence 24, Appli
28	57.5	55.3	38	12	US-10-325-810-214	Sequence 21, App
29	49	47.1	2169	9	US-09-738-626-5455	Sequence 5455, Ap
30	48.5	46.6	38	9	US-09-843-676-25	Sequence 25, Appli
31	48.5	46.6	38	10	US-09-766-253-25	Sequence 25, Appli
32	48.5	46.6	38	10	US-09-438-486-25	Sequence 25, Appli
33	48.5	46.6	38	12	US-10-325-810-215	Sequence 215, App
34	48.5	46.6	38	14	US-10-053-758-25	Sequence 25, Appli
35	48.5	46.6	38	14	US-10-054-295-25	Sequence 25, Appli
36	48.5	46.6	38	14	US-10-054-611-25	Sequence 25, Appli
37	47	45.2	456	9	US-09-861-400-6	Sequence 6, Appli
38	47	45.2	482	9	US-09-861-400-5	Sequence 5, Appli
39	46.5	44.7	16	9	US-09-836-073-19	Sequence 19, Appli
40	45.5	43.8	425	12	US-10-296-115-1146	Sequence 1146, Ap
41	45.5	43.8	643	9	US-09-736-968A-60	Sequence 60, Appli
42	45.5	43.8	675	10	US-09-978-244A-18	Sequence 18, Appli
43	45.5	43.8	2045	9	US-09-736-969A-95	Sequence 95, Appli
44	45.5	43.8	2045	9	US-09-736-960-92	Sequence 92, Appli
45	45.5	43.8	2045	9	US-09-736-968A-109	Sequence 109, App

ALIGNMENTS

RESULT 1
US-09-836-073-16
; Sequence 16, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Rat
US-09-836-073-16

Query Match 100.0%; Score 104; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.2e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEHYFGDF 19
DB 1 AALEAKICHQIEHYFGDF 19

RESULT 2
US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.

Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 7 ICHQIEBYYPGDF 19
| | | | | | | | | |
Db 147 IAHQIEGYFMGHF 159

Search completed: September 10, 2004, 18:00:12
Job time : 46.5196 secs

Best Local Similarity 56.2%; Pred. No. 21;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 4 EAARICHOIEEYFGDF 19
Db 128 ELKFNHIVNEFSFGDF 143

RESULT 13

ID 029131 PRELIMINARY; PRT; 456 AA.

AC 029131; 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
GN AF1134.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
Ketchum K.A., Dodson R.J., Gilm M., Hickey E.K., Peterson J.D.,
Richardson D.L., Krelavage A.R., Graham D.E., Kyriades N.C.,
R. Kirchmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodok A., Zhou L.,
Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
Cotton M.D., Spriggs T., Arriach P., Kaine B.P., Sykes S.M.,
Sadov P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
Venter J.C.;
RA "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus.",
RT Nature 390:364-370(1997).
RL EMBL, A8001026; AAB90123.1; -.
DR PIR; E69391; E69391.
DR TIGR; AF1134; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR000379; Ser_easts.
KW Hypothetical protein, Complete proteome.
SQ SEQUENCE 456 AA; 51748 MW; 37CB62B58C2C9357 CRC64;

Query Match 44.2%; Score 46; DB 17; Length 456;
Best Local Similarity 61.5%; Pred. No. 35;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 LEAKICHOIEEY 15
Db 405 MEAKTPHOFDEY 417

RESULT 14

ID 0727P4 PRELIMINARY; PRT; 509 AA.

AC 0727P4; 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22388257; PubMed=12477932;
RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Scheefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Franke C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettlemen M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA "Trzywnicki M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Straube R.L.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL, BC051330; AAH51330.1; -.
FT NON_TER 1
SQ SEQUENCE 509 AA; 58441 MW; CD75C36F3BF8580 CRC64;

Query Match 43.8%; Score 45.5; DB 4; Length 509;
Best Local Similarity 52.9%; Pred. No. 48;
Matches 9; Conservative 4; Mismatches 1; Indels 3; Gaps 1;

QY 5 AKICHOIEEY---FGD 18
Db 242 AETSHRLSEYTERFGD 258

RESULT 15

ID 08GK02 PRELIMINARY; PRT; 166 AA.

AC 08GK02; 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
GN Glutamate racemase (Fragment).
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PS149A1;
RA Dallide G., Ogura M., Dallidene D., Berg D.E.,
RT "Geographic partitioning in Helicobacter pylori: Gene pools of Spain and
Peru are closely related, and distinct from those of Japan.";
RT Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL, AY15315; AAN74490.1; -.
DR 30; GO:0008881; F:glutamate racemase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001920; Asp/Glu_race.
DR InterPro; IPR004391; Glu_race.
DR Pfam; PF01177; Asp_Glu_race; 1.
DR TIGRFAMs; TIGR00067; glu_race; 1.
DR PROSITE; PS00923; ASP_GLU_RACEMASE_1; 1.
DR PROSITE; PS00924; ASP_GLU_RACEMASE_2; 1.
FT NON_TER 1
FT NON_TER 166
SQ SEQUENCE 166 AA; 18429 MW; 323D0AD415B6185 CRC64;

Query Match 43.3%; Score 45; DB 2; Length 166;
Best Local Similarity 61.5%; Pred. No. 18;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHQIEEYFGD 18
|||:|:|
Db 97 ICHELESWFLGD 108

RESULT 9

QEX37 PRELIMINARY; PRT; 301 AA.
ID QEX37;
AC QEX37;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Pseudouridine synthase.
GN MYR130.
OS Mycoplasma penetrans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
NX NCBI_TaxID=28227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HF-2;
RA MEDLINE=22354719; PubMed=1246555;
RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.,
RT "The complete genomic sequence of Mycoplasma penetrans, an
RT intracellular bacterial pathogen in humans."
RL Nucleic Acids Res. 30:5293-5300 (2002).
DR EMBL, AP004170; BAC43803.1; -
DR GO: GO:0004730; F:pseudouridylylate synthase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR006145; Pseudou_synth.
DR InterPro; IPR006225; Pseudou_synth.
DR InterPro; IPR006224; Rlu_synth.
DR Pfam; PF00849; Pseudou_synth_2; 1.
DR ProDom; PD001819; PSI_RLU; 1.
DR TIGRFAMs; TIGR00005; TUD_subfam; 1.
DR PROSITE; PS50889; S4; 1.
KM Complete proteome.
SQ SEQUENCE 301 AA; 35770 MW; 2961D557AB45B730 CRC64;

Query Match 45.2%; Score 47; DB 16; Length 301;
Best Local Similarity 46.7%; Pred. No. 15;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 LEAKICHOIEEYFG 17
|||:|:|
Db 131 LEANICHRUDKXTTG 145

RESULT 10

Q8SC2 PRELIMINARY; PRT; 357 AA.
ID Q8SC2;
AC Q8SC2;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein orf357 from chromosome 3.
GN ORF357.
OS Guillardia theta (Cryptomonas phi).
OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
NX NCBI_TaxID=55529;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2123349; PubMed=1123671;
RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.T.,
RA Wu X., Reich M., Cavalier-Smith T., Maier U.G.,
RT "The highly reduced genome of an enslaved algal nucleus."
RL Nature 410:1091-1096 (2001).
DR EMBL, AF083031; AAK39661.1; -
DR PIR; H90120; H90120.
DR InterPro; IPR001680; WD40.

DR SMART; SM00320; WD40; 3.
KM Hypothetical protein; Repeat; WD repeat
SQ SEQUENCE 357 AA; 42401 MW; 71AB9F896A04C82 CRC64;

Query Match 45.2%; Score 47; DB 10; Length 357;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 LEAKICHOIEEYFGD 18
|||:|:|
Db 242 LKNVFSQMKRYFGD 257

RESULT 11

Q42907 PRELIMINARY; PRT; 448 AA.
ID Q42907;
AC Q42907;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN SPBC119.16C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
NX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97h-;
RA Wood V., Rajandream M.A., Barrell B.G., Skelton J., Churcher C.M.,
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL, AL022117; CAI17931.1; -
DR PIR; T39314; T39314.
DR GenDB; SPombe; SPBC119.16C; -
KM Hypothetical protein.
SQ SEQUENCE 448 AA; 51190 MW; 3B56729FA1A98DAB CRC64;

Query Match 45.2%; Score 47; DB 3; Length 448;
Best Local Similarity 47.4%; Pred. No. 23;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 AALEAKICHOIEEYFGD 19
|||:|:|
Db 9 ATPDAKEGIVNENYFGD 27

RESULT 12

Q04639 PRELIMINARY; PRT; 284 AA.
ID Q04639;
AC Q04639;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Chromosome XIII COSMID 9745.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB972;
RA Bowman S., Churcher C.,
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB972;
RA Barrell B., Rajandream M.A.,
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z38114; CAA86253.1; -
SQ SEQUENCE 284 AA; 32882 MW; 200EB06A92A3CCF CRC64;

Query Match 44.2%; Score 46; DB 3; Length 284;

DR GO:0005634; C:nucleus; IDA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF05383; LA; 1.
DR Pfam; PF00076; xrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
SQ SEQUENCE 381 AA; 43891 MW; 2E2DEF1452C0F0E9 CRC64;

Query Match 81.2%; Score 84.5; DB 11; Length 381;
Best Local Similarity 94.4%; Pred. No. 1.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 2 ALEAKICHOIEEYFGDF 19
Db 12 ALEAKICHOI-EYFGDF 28

RESULT 2

Q8BTU4 PRELIMINARY; PRT; 415 AA.
AC Q8BTU4;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Sjogren syndrome antigen B.
GN SSB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOB; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK088677; BAC40498.1; -.
DR MGD; MGI:98423; Seb.
DR GO; GO:0005634; C:nucleus; IDA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF05383; LA; 1.
DR Pfam; PF00076; xrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 415 AA; 47657 MW; A7545C7686AC8363 CRC64;

Query Match 81.2%; Score 84.5; DB 11; Length 415;
Best Local Similarity 94.4%; Pred. No. 1.2e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 2 ALEAKICHOIEEYFGDF 19
Db 12 ALEAKICHOI-EYFGDF 28

RESULT 3

Q7ZTK2 PRELIMINARY; PRT; 427 AA.
AC Q7ZTK2;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Similar to lupus LA protein homolog B.

OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenoipodinae; Xenopus.
OX NCBI_TaxId=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC046654; AA046654.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00076; xrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 427 AA; 48996 MW; 1E7CD82D8AB9C69A CRC64;

Query Match 57.2%; Score 59.5; DB 13; Length 427;
Best Local Similarity 75.0%; Pred. No. 0.19;
Matches 12; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

Qy 3 LEAKICHOIEEYFGD 18
Db 12 LDTKICEGI-EYFGD 26

RESULT 4

Q8OHI5 PRELIMINARY; PRT; 206 AA.
AC Q8OHI5;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Untranslated region binding-protein.
GN UBP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxId=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA L'Ecuylar T.J., Fang H.-L.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF467897; AAL76269.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF05383; LA; 1.
DR Pfam; PF00076; xrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 206 AA; 23992 MW; 965B62FDFFB90E9 CRC64;

Query Match 56.2%; Score 58.5; DB 13; Length 206;
Best Local Similarity 75.0%; Pred. No. 0.13;
Matches 12; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 3 LEAKICHOIEEYFGD 18
Db 13 LESKICQOI-EYFGN 27

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:29:01; Search time 43.5196 Seconds
(without alignments)
137.751 Million cell updates/sec

Title: US-09-836-073-16

Perfect score: 104
Sequence: 1 AALEAKICHQIEHYFGDF 19

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL, 25: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriophage: *
17: sp_archaeal: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84.5	81.2	381	11	Q9CYB9 mus musculus
2	84.5	81.2	415	11	Q8BTU4 mus musculus
3	59.5	57.2	427	13	Q7ZTK2 xenopus lae
4	58.5	56.2	266	13	Q8OHIS gallus gall
5	55	52.9	135	16	Q67676 aquifex aeo
6	51.5	49.5	401	13	Q7ZT10 brachydanio
7	49	47.1	2169	16	Q8NPN6 corynebacte
8	47	45.2	187	16	Q8YF2 anabaena sp
9	47	45.2	301	16	Q8YF2 anabaena sp
10	47	45.2	357	10	Q98SC2 guillardia
11	47	44.2	448	3	Q42907 schizosacch
12	46	44.2	284	3	Q04639 saccharomyc
13	46	44.2	456	17	Q29131 archaeoglob
14	45.5	43.8	404	7	Q7Z7P4 homo sapien
15	45	43.3	166	2	Q8GK02 helicobacte
16	45	43.3	184	2	Q7X5D8 helicobacte

17	45	43.3	231	2	Q7X5E0	Q7X5E0 helicobacte
18	45	43.3	409	8	Q98P4	Q98P4 guillardia
19	45	43.3	748	16	Q8AX3	Q8AX3 bacteroides
20	45	43.3	928	10	Q9LJ02	Q9LJ02 oryza sativ
21	45	43.3	1156	12	Q57230	Q57230 vaccinia vi
22	45	43.3	1164	12	Q90027	Q90027 variola maj
23	45	43.3	1164	12	Q9JF79	Q9JF79 vaccinia vi
24	45	43.3	1164	12	Q9JF79	Q9JF79 vaccinia vi
25	45	43.3	1164	12	Q8L190	Q8L190 ectromelia
26	45	43.3	1164	12	Q8V4V3	Q8V4V3 monkeypox v
27	45	43.3	1164	12	Q8V2N1	Q8V2N1 camelpox vi
28	45	43.3	1164	12	Q8DVL1	Q8DVL1 camelpox vi
29	45	43.3	1676	8	Q85FR6	Q85FR6 cyanidiosch
30	44	42.3	281	11	Q9COK3	Q9COK3 mus musculu
31	44	42.3	284	4	Q8NG50	Q8NG50 homo sapien
32	44	42.3	505	8	Q85V94	Q85V94 elaeagnu s
33	44	42.3	592	5	Q9VN01	Q9VN01 drosophila
34	44	42.3	627	5	Q81EX1	Q81EX1 trypanosoma
35	44	42.3	627	5	Q86MC2	Q86MC2 trypanosoma
36	44	42.3	627	5	Q86MC2	Q86MC2 trypanosoma
37	44	42.3	658	5	Q86B82	Q86B82 drosophila
38	44	42.3	888	3	Q873B9	Q873B9 neurospora
39	43.5	41.8	119	10	Q9ZP89	Q9ZP89 neurotheca
40	43	41.3	185	5	Q9UAF3	Q9UAF3 polyantheca
41	43	41.3	297	4	Q9H1U9	Q9H1U9 homo sapien
42	43	41.3	427	16	Q05874	Q05874 mycobacteri
43	43	41.3	427	16	Q7TWY2	Q7TWY2 mycobacteri
44	43	41.3	447	10	Q93V61	Q93V61 arabidopsis
45	43	41.3	510	10	Q43819	Q43819 plasmu sativ

ALIGNMENTS

RESULT 1
ID Q9CYB9 PRELIMINARY; PRT; 381 AA.
AC Q9CYB9; 01-JUN-2001 (T-REMBLrel. 17, Created)
DT 01-JUN-2001 (T-REMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (T-REMBLrel. 25, Last annotation update)
DE Sjogren syndrome antigen B.
OS SSB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=1117051;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana S.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giesi C., King B., Kociwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staudli F., Suzuki R., Tomita M., Wagner L., Maehio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Bartsch G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita K., Gariboldi M.,
RA Gusttinch S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Maehima J., Mazzarelli U., Mombarts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki K., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitlaker C., Wilming L.,
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK017822; BAB30957.1; -.
DR MGD; MGI:98423; Ssb.

DR Pfam; PF00562; RNA_pol_Rpb2_6; 1.
 DR Pfam; PF04560; RNA_pol_Rpb2_7; 1.
 DR PROSITE; PS01166; RNA_POL_BETA; 1.
 KW Transferase; DNA-directed RNA polymerase; transcription; zinc;
 KW zinc-finger.
 FT ZN PING 1087 1106 C4-TYPE (POTENTIAL).
 SQ SEQUENCE 1164 AA; 133329 MW; 0452B84ED810CD53 CRC64;

Query Match 43.3%; Score 45; DB 1; Length 1164;
 Best local Similarity 56.2%; Pred. NO. 22;
 Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 3 LEAKICHQIEHYFGD 18
 Db 505 LEKKICEYIRSYKDD 520

Search completed: September 10, 2004, 17:53:09
 Job time : 9.64246 secs

FT INIT MET 9 9 POLYPEPTIDE, ISOFORM 2.
 FT ZN FING 1087 1106 C4-TYPE (POTENTIAL).
 SQ SEQUENCE 1164 AA; 133363 MW; 220DBE5F5238DB7 CRC64;
 Query Match 43.3%; Score 45; DB 1; Length 1164;
 Best Local Similarity 56.2%; Pred. No. 22;
 Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 3 LEAKICHOIEEYFGD 18
 DB 505 LEKICIEYIRSYKDD 520

RESULT 14
 ID RPO2 VACCV STANDARD; PRT; 1164 AA.
 AC P19798;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA-directed RNA polymerase 132 kDa polypeptide (EC 2.7.7.6).
 GN RPO132 OR A24R.
 OS Vaccinia virus (strain WR), and
 OS Vaccinia virus (strain Copenhagen).
 OS Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 NCBI_TaxID=10254, 10249;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WR;
 RX MEDLINE=91082452; PubMed=1824607;
 RA Amegadze B.Y., Holmes M.H., Cole N.B., Jones E.V., Earl P.L.,
 RA Moss B.;
 RT "Identification, sequence, and expression of the gene encoding the
 RT second-largest subunit of the vaccinia virus DNA-dependent RNA
 RT polymerase.";
 RL Virology 180:88-96(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Copenhagen;
 RX MEDLINE=91021027; PubMed=2219722;
 RA Goebel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,
 RA Paolletti E.;
 RT "The complete DNA sequence of vaccinia virus.";
 RL Virology 179:247-266(1990).
 RN [3]
 RP COMPLETE GENOME.
 RC STRAIN=Copenhagen;
 RA Goebel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,
 RA Paolletti E.;
 RT "Appendix to 'The complete DNA sequence of vaccinia virus.'";
 RL Virology 179:517-563(1990).
 CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
 CC of DNA into RNA using the four ribonucleoside triphosphates as
 CC substrates.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}(N).
 CC -1- SUBUNIT: This enzyme consists of at least eight subunits.
 CC -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M37415; AAA72882.1; -
 CC EMBL; M35027; AAA48148.1; -
 CC PIR; H42519; RNZ8T.
 CC InterPro; IPR007121; RNA_pol_B.

DR InterPro; IPR007645; RNA_pol_Rpb2_3.
 DR InterPro; IPR007646; RNA_pol_Rpb2_4.
 DR InterPro; IPR007647; RNA_pol_Rpb2_5.
 DR InterPro; IPR007120; RNA_pol_Rpb2_6.
 DR InterPro; IPR007641; RNA_pol_Rpb2_7.
 DR Pfam; PF04565; RNA_pol_Rpb2_3; 1.
 DR Pfam; PF04566; RNA_pol_Rpb2_4; 1.
 DR Pfam; PF04567; RNA_pol_Rpb2_5; 1.
 DR Pfam; PF04568; RNA_pol_Rpb2_6; 1.
 DR Pfam; PF04569; RNA_pol_Rpb2_7; 1.
 DR PROSITE; PS01166; RNA_Pol_BETA.1.
 DR Transferrase; DNA-directed RNA polymerase; Transcription; Zinc;
 KW Zinc-finger.
 FT ZN FING 1087 1106 C4-TYPE (POTENTIAL).
 SQ SEQUENCE 1164 AA; 133363 MW; B1F5C1484BA37D0D CRC64;
 Query Match 43.3%; Score 45; DB 1; Length 1164;
 Best Local Similarity 56.2%; Pred. No. 22;
 Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 3 LEAKICHOIEEYFGD 18
 DB 505 LEKICIEYIRSYKDD 520

RESULT 15
 ID RPO2 VARV STANDARD; PRT; 1164 AA.
 AC P33811;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE DNA-directed RNA polymerase 132 kDa polypeptide (EC 2.7.7.6).
 GN RPO132 OR A24R.
 OS Variola virus.
 OS Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 NCBI_TaxID=10255;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=India-1967 / Isolate Ind3;
 RX MEDLINE=93202281; PubMed=8384129;
 RA Shchelkunov S.N., Blinov V.M., Sandakhchiev L.S.;
 RT "Genes of variola and vaccinia viruses necessary to overcome the host
 RT protective mechanisms.";
 RL FEBS Lett. 319:80-83(1993).
 CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
 CC of DNA into RNA using the four ribonucleoside triphosphates as
 CC substrates.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}(N).
 CC -1- SUBUNIT: This enzyme consists of at least eight subunits.
 CC -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X69198; CAA49069.1; -
 CC PIR; G36850; G36850.
 CC InterPro; IPR007121; RNA_pol_B.
 CC InterPro; IPR007645; RNA_pol_Rpb2_3.
 CC InterPro; IPR007646; RNA_pol_Rpb2_4.
 CC InterPro; IPR007647; RNA_pol_Rpb2_5.
 CC InterPro; IPR007120; RNA_pol_Rpb2_6.
 CC InterPro; IPR007641; RNA_pol_Rpb2_7.
 CC Pfam; PF04565; RNA_pol_Rpb2_3; 1.
 CC Pfam; PF04566; RNA_pol_Rpb2_4; 1.
 CC Pfam; PF04567; RNA_pol_Rpb2_5; 1.
 CC Pfam; PF04568; RNA_pol_Rpb2_6; 1.
 CC Pfam; PF04569; RNA_pol_Rpb2_7; 1.

AC P38433: 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Acetylcholinesterase 1 precursor (EC 3.1.1.7) (ACNE 1).
 GN ACN-1 OR W09812.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=94193691; PubMed=8144590;
 RA Arpagaus M., Fedon Y., Cousin X., Chatonnet A., Berge J.-B.,
 RA Fournier D., Toulant J.-P.;
 RT "cDNA sequence, gene structure, and in vitro expression of ace-1, the
 RT gene encoding acetylcholinesterase of class A in the nematode
 RT Caenorhabditis elegans.";
 RL J. Biol. Chem. 269:9957-9965 (1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Mu X., Le T.T.;
 RU Submitted (May-1996) to the EMBL/GenBank/DBSJ databases.
 CC -1- FUNCTION: RAPIDLY HYDROLYZES CHOLINE RELEASED INTO THE SYNAPSE.
 CC IT CAN HYDROLYZE BUTYRYLTHIOCHOLINE.
 CC -1- CATALYTIC ACTIVITY: Acetylcholine + H(2)O = choline + acetate.
 CC -1- SUBUNIT: Oligomer composed of disulfide-linked homodimers (by
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: MAY BE SECRETED OR MEMBRANE ASSOCIATED VIA
 CC A NON-CATALYTIC SUBUNIT.
 CC -1- DEVELOPMENTAL STAGE: DETECTED AT ALL STAGES. FOUND TO BE MORE
 CC ABUNDANT IN LARVAL STAGES THAN IN EMBRYOS OR ADULTS.
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X7531; CAA53080.1; -;
 DR EMBL: U58731; AAB00593.1; -;
 DR PIR: A54413; A54413.
 DR HSSP: P21836; 1MAA.
 DR WormRep: W09812.1; CE07569.
 DR InterPro: IPR002018; Carboxylesterase.
 DR InterPro: IPR000997; Cholinesterase.
 DR InterPro: IPR000379; Ser esters.
 DR Pfam: PF00135; Coesterase; 1.
 DR PRINTS: PR00878; CHOLINESTERASE.
 DR PROSITE: PS00122; CARBOXYL-ESTERASE B_1; 1.
 DR PROSITE: PS00941; CARBOXYL-ESTERASE B_2; 1.
 DR Hydrolase; Serine esterase; Synapse; Membrane; Nerve; Signal;
 DR Neutransmitter degradation; Multigene family.
 DR SIGNAL 1
 FT CHAIN 1 31
 FT ACT_SITE 32 620 ACETYLCHOLINESTERASE 1.
 FT ACT_SITE 216 216
 FT ACT_SITE 346 346 BY SIMILARITY.
 FT ACT_SITE 468 468 BY SIMILARITY.
 FT DISULFID 82 109 BY SIMILARITY.
 FT DISULFID 270 286 BY SIMILARITY.
 FT DISULFID 430 558 BY SIMILARITY.
 FT DISULFID 618 618 INTERCHAIN (BY SIMILARITY).
 FT CARBOXHD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOXHD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOXHD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOXHD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SEQUENCE 620 AA; 71433 MW; 61D78C4899F55C65 CRC64;

Query March 43.3%; Score 45; DB 1; Length 620;
 Best Local Similarity 66.7%; Pred. No. 11;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Qy 8 CHOIEEYFGDF 19
 Db 82 CIOSEDTYFGDF 93
 RESULT 13
 RPO2 COMPLEX STANDARD; PRT; 1164 AA.
 AC P17474; Q90025;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE DNA-directed RNA polymerase 132 kDa polypeptide (EC 2.7.7.6).
 GN RPO132.
 OS Compo virus (CPV).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OX NCBI_TaxID=10243;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Brichon red;
 RX MEDLINE=89125698; PubMed=2915377;
 RA Patel D.D., Pickup D.V.;
 RT "The second-largest subunit of the poxvirus RNA polymerase is similar
 RT to the corresponding subunits of procaryotic and eucaryotic RNA
 RT polymerases.";
 RL J. Virol. 63:1076-1086 (1989).
 CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
 CC of DNA into RNA using the four ribonucleoside triphosphates as
 CC substrates.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA} (N).
 CC -1- SUBUNIT: This enzyme consists of at least eight subunits.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=alternative initiation;
 CC Comment=2 isoforms, 1/late specific (shown here) and 2/early and
 CC late, are produced by alternative initiation. One
 CC transcriptional start site is operative at late times only and
 CC the other is operative both at early and late times;
 CC -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M26173; AAA42919.1; -;
 DR EMBL: M26173; AAA42920.1; -;
 DR PIR: A31879; RNVCZP.
 DR InterPro: IPR007121; RNA_pol_B.
 DR InterPro: IPR007645; RNA_pol_Rpb2_3.
 DR InterPro: IPR007646; RNA_pol_Rpb2_4.
 DR InterPro: IPR007647; RNA_pol_Rpb2_5.
 DR InterPro: IPR007120; RNA_pol_Rpb2_6.
 DR InterPro: IPR007641; RNA_pol_Rpb2_7.
 DR Pfam: PF04565; RNA_pol_Rpb2_3; 1.
 DR Pfam: PF04566; RNA_pol_Rpb2_4; 1.
 DR Pfam: PF04567; RNA_pol_Rpb2_5; 1.
 DR Pfam: PF00562; RNA_pol_Rpb2_6; 1.
 DR Pfam: PF04560; RNA_pol_Rpb2_7; 1.
 DR PROSITE: PS01166; RNA_POL_BETA_1.
 DR Transferase; DNA-directed RNA polymerase; Transcription; Zinc;
 DR Zinc-finger; Alternative initiation.
 DR CHAIN 1 1164
 FT CHAIN 9 1164
 FT SEQUENCE 1164 AA; 12644 MW; 61D78C4899F55C65 CRC64;
 FT CHAIN 9 1164
 FT SEQUENCE 1164 AA; 12644 MW; 61D78C4899F55C65 CRC64;

RT "Construction of expression-ready cDNA clones for KIAA genes: manual
RT curation of 330 KIAA cDNA clones."
RL DNA Res. 9:99-106(2002).
RN [3]
RP SEQUENCE OF 1188-2047 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schler G.D.,
RA Altchul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mllahy S.J.,
RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richard S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP NOMENCLATURE.
RX MEDLINE=22319137; PubMed=12432077;
RA Cote J.-F., Vuori K.;
RT "Identification of an evolutionarily conserved superfamily of DOCK180-
RT related proteins with guanine nucleotide exchange activity."
RL J. Cell Sci. 115:4901-4913(2002).
CC -1- FUNCTION: Potential guanine nucleotide exchange factor (GEF). GEF
CC proteins activate some small GTPases by exchanging bound GDP for
CC free GTP (By similarity).
CC -1- TISSUE SPECIFICITY: Widely expressed. Expressed at low level in
CC spleen, cerebellum, hippocampus and in substantia nigra.
CC -1- DOMAIN: The DHR-2 domain may mediate some GEF activity (By
CC similarity).
CC -1- SIMILARITY: Belongs to the DOCK family.
CC -1- SIMILARITY: Contains 1 DHR-1 (CZH-1) domain.
CC -1- SIMILARITY: Contains 1 DHR-2 (CZH-2) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB037816; BAA92633.2; ALT_INIT.
CC EMBL; BC008335; AA08335.1; -.
CC Genew; HGNC:19189; DOCK6.
CC InterPro; IPR002016; Peroxidase.
CC Guanine-nucleotide releasing factor.
CC FT DOMAIN 548 786 DHR-1.
CC FT DOMAIN 1489 2012 DHR-2.
CC SO SEQUENCE 2047 AA; 229655 MW; 55F4BDDAB1E60F31 CRC64;
QY Query Match 43.8%; Score 45.5; DB 1; Length 2047;
Best Local Similarity 52.9%; Pred. No. 33;
Matches 9; Conservative 4; Mismatches 1; Indels 3; Gaps 1;
Db 5 AKICHOIEEY---FGD 18
1780 AETSHRLREYTERFGD 1796
RESULT 11
ACE1_CAEBR

ID ACE1_CAEBR STANDARD; PRT; 620 AA.
AC 027459;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acetylcholinesterase 1 precursor (EC 3.1.1.7) (ACHe 1).
GN ACE-1.
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_Taxid=6238;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97069944; PubMed=8912924;
RA Grauso M., Culetto E., Borge J.B., Toulant J.-P., Arpagaus M.;
RT "Sequence comparison of ACE-1, the gene encoding acetylcholinesterase
RT of class A, in the two nematodes Caenorhabditis elegans and
RT Caenorhabditis briggsae."
RL DNA Seq. 6:217-227(1996).
CC -1- FUNCTION: Rapidly hydrolyzes choline released into the synapse.
CC -1- CATALYTIC ACTIVITY: Acetylcholine + H(2)O = choline + acetate.
CC -1- SUBUNIT: oligomer composed of disulfide-linked homodimers (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: MAY BE SECRETED OR MEMBRANE ASSOCIATED VIA
CC A NON-CATALYTIC SUBUNIT (BY SIMILARITY).
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC HMBL; U41846; AAB41269.1; -.
CC HSSP; P21836; IMAA.
CC InterPro; IPR002018; Carboxylesterase.
CC InterPro; IPR000997; Cholinesterase.
CC InterPro; IPR000379; Ser_estrs.
CC Pfam; PF00135; Coesterase_1.
CC PRINTS; PR00878; CHOLINESTERASE.
CC PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
CC PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
CC KW Hydrolyase; Serine esterase; Synapse; Membrane; Signal; Glycoprotein;
CC Neurotransmitter degradation; Multigene family.
CC -----
CC FT SIGNAL 1 31 POTENTIAL.
CC FT CHAIN 32 620 ACETYLCHOLINESTERASE 1.
CC FT ACT SITE 216 216 BY SIMILARITY.
CC FT ACT SITE 346 346 BY SIMILARITY.
CC FT ACT SITE 468 468 BY SIMILARITY.
CC FT DISULFID 82 109 BY SIMILARITY.
CC FT DISULFID 270 286 BY SIMILARITY.
CC FT DISULFID 430 558 BY SIMILARITY.
CC FT DISULFID 618 618 INTERCHAIN (BY SIMILARITY).
CC FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 486 486 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SO SEQUENCE 620 AA; 71501 MW; 69D73CD3996E11FC CRC64;
QY Query Match 43.3%; Score 45; DB 1; Length 620;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Db 8 CHQIEEYFGDF 19
82 CISEDYFGDF 93
RESULT 12
ACE1_CAEBL STANDARD; PRT; 620 AA.
ID ACE1_CAEBL

KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 24 44 POTENTIAL.
 FT TRANSMEM 290 310 POTENTIAL.
 FT TRANSMEM 312 332 POTENTIAL.
 SQ SEQUENCE 352 AA; 40705 MW; A9P02FEB9766501 CRC64;
 Query Match 44.2%; Score 46; DB 1; Length 352;
 Best Local Similarity 56.2%; Pred No. 4.2;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 4 BAKICHOIEEYFGDF 19
 DB 196 ELKFMHVINFEFGDF 211
 RESULT 9
 DOC6_MOUSE STANDARD; PRT; 849 AA.
 AC 08VDR9; Q9D461;
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dedicator of cytokinesis protein 6 (Fragment).
 DOCK6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain, and Breast tumor;
 RX MEDLINE=22386257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein W.J., Uebli B., Toshyuki S., Carninci P., Prange C.,
 Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Millaby S.J.,
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultik S.W.,
 Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
 Fahy J., Helton E., Kerteman M., Madan A., Rodrigues S., Sanchez A.,
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE OF 640-849 FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=22554683; PubMed=12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 Nishida I., Oatso N., Saito R., Suzuki H., Yamahata I., Kiyosawa H.,
 Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,
 Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 Schiml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 Blake J., Brad T., Brusic V., Chotia C., Corbani L.E., Cousins S.,
 Dalla E., Dragan T., Fletcher C.F., Forrest A., Frazer K.S.,
 Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 Grimond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 Kani A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 Kongsava A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 Nagashima T., Numa K., Okido T., Pavan W.J., Petrea G., Pesole G.,
 Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 Sandelin A., Schneider C., Sempé C.A., Setou M., Shimada K.,
 Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,

RA Verardo R., Wagner L., Wanhstedt C., Wang Y., Watanabe Y., Wells C.,
 Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazune N., Sato K.,
 Shitaka T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
 Yaunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 Birney E., Hayashizaki Y.,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 CC -1- FUNCTION: Potential guanine nucleotide exchange factor (GEF). GEF
 proteins activate some small GTPases by exchanging bound GDP for
 free GTP (By similarity).
 CC -1- DOMAIN: The DHR-2 domain may mediate some GEF activity (By
 similarity).
 CC -1- SIMILARITY: Belongs to the DOCK family.
 CC -1- SIMILARITY: Contains 1 DHR-2 (CZH-2) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 DR EMBL, BC021414; AAH21414.1; ALT_INIT.
 DR EMBL, BC043042; AAH43042.1; -.
 DR EMBL, AK016777; BAB30423.2; -.
 DR MGI:1914789; 4931431C02R1.
 KW Guanine-nucleotide releasing factor.
 FT NON_TER 1
 FT DOMAIN 1
 SQ SEQUENCE 849 AA; 95975 MW; 5374B30CB97265CE CRC64;
 QY 5 AKICHOIEEY---FGD 18
 DB 582 A6ISHRLFEFTTERGD 598
 RESULT 10
 DOC6_HUMAN STANDARD; PRT; 2047 AA.
 AC 096H0; Q9P2E2;
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dedicator of cytokinesis protein 6.
 DOCK6 OR KIAA1395.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC TISSUE=Brain;
 RX MEDLINE=2018126; PubMed=10718198;
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirogawa M., Ohara O.,
 RT "Prediction of the coding sequences of unidentified human genes. XVI.
 RT The complete sequences of 150 new cDNA clones from brain which code
 RT for large proteins in vitro."
 RL DNA Res. 7:65-73(2000).
 RN [2]
 RP REVISIONS.
 RC TISSUE=Brain;
 RX MEDLINE=22158633; PubMed=12168954;
 RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;

```

CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X68817; CAA48715.1; '-'
CC PIR; S33818; S33818.
CC InterPro; IPR002344; Lupus_La.
CC InterPro; IPR006630; Lupus_La_dom.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF05383; La; 1.
CC Pfam; PF00076; xrm; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00715; LA; 1.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS50102; RRM; 1.
CC PROSITE; PS00030; RRM_RNP_1; 1.
CC RNA-binding; Nuclear protein; Phosphorylation.
CC DOMAIN 111 203 RNA-BINDING (RRM).
CC DOMAIN 166 212 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC DOMAIN 316 332 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC SEQUENCE 428 AA; 48664 MW; AEB3A5B7D2B3EC3 CRC64;
CC -----
CC Query Match 53.4%; Score 55.5; DB 1; Length 428;
CC Best Local Similarity 73.3%; Pred. No. 0.14;
CC Matches 11; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
CC -----
CC 4 EAKTCHQIEEYFGD 18
CC : ||| |||||
CC 14 DTKICEQI-EYVFGD 27
CC -----
CC RESULT 7
CC LBP_RABIT STANDARD; PRT; 482 AA.
CC ID LBP RABIT
CC AC P17454;
CC DT 01-AUG-1990 (Rel. 15, Created)
CC DT 01-NOV-1990 (Rel. 16, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Lipopolysaccharide-binding protein precursor (LBP).
CC GN LBP.
CC OS Oryctolagus cuniculus (Rabbit).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Butheria; Lagomorpha; Leporidae; Oryctolagus.
CC OX NCBI_TaxID=9986;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=90385281; PubMed=2402637;
CC RA Schumann R.R., Leong S.R., Flagg G.W., Gray P.W., Wright S.D.,
CC Mathison J.C., Tobias P.S., Ulevitch R.J.;
CC "Structure and function of lipopolysaccharide binding protein.";
CC Science 249:1429-1431(1990).
CC RN [2]
CC RP SEQUENCE OF 27-66.
CC RC TISSUE-Serum;
CC RX MEDLINE=86306528; PubMed=2427635;
CC RA Tobias P.S., Soldau K., Ulevitch R.J.;
CC RT "Isolation of a lipopolysaccharide-binding acute phase reactant from
CC rabbit serum.";
CC RU J. Exp. Med. 164:777-793(1986).
CC CC -I- FUNCTION: Binds to the lipid moiety of bacterial
CC lipopolysaccharides (LPS), a glycolipid present in the outer
CC membrane of all Gram-negative bacteria. The LBP/LPS complex seems
CC to interact with the CD14 receptor.
CC CC -I- SUBCELLULAR LOCATION: Secreted.
CC CC -I- SIMILARITY: Belongs to the BPI/LBP/plunc superfamily. BPI/LBP
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC      or send an email to license@lsb-sib.ch.
CC      -----
DR      EMBL, M35534; AA99235.1; -.
DR      PIR, B35843; B35843.
DR      HSSP, P17213; 1BP1.
DR      InterPro, IPR001124; LBP_BPI_CETP.
DR      Pfam, PF01273; LBP_BPI_CETP, 1.
DR      Pfam, PF02886; LBP_BPI_CETP_C, 1.
DR      SMART, SM00328; BPI1; 1.
DR      SMART, SM00329; BPI2; 1.
DR      ProSite, PS00400; LBP_BPI_CETP, 1.
KW      lipid transport; Antibiotic; Transmembrane; Glycoprotein; Signal.
FT      CHAIN 1 26
FT      CARBOYD 301 301 LIPOPOLYSACCHARIDE-BINDING PROTEIN.
FT      CARBOYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CONFLICT 57 57 E -> G (IN REF. 2).
FT      CONFLICT 63 63 S -> F (IN REF. 2).
SQ      SEQUENCE 482 AA; 54001 MW; 628AE0A647200C2 CRC64;
OY      Query Match 45.2%; Score 47; DB 1; Length 482;
        Best Local Similarity 81.8%; Pred. No. 4;
        Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY      3 LEAKCHOICE 13
Db      194 LESKICROIEB 204

RESULT 8
YMG7_YEAST STANDARD; PRT; 352 AA.
AC      Q04651;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Hypothetical 40.7 kDa protein in DAK1-ORC1 intergenic region.
GN      YML067C.
OC      Saccharomyces cerevisiae (Baker's yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Saccharomyces.
CX      NCBI_TaxID=4932;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=5288C / AB972;
RX      MEDLINE=97313268; PubMed=9169872;
RA      Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA      Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA      Jagels K., Iye G., Moule S., Odell C., Pearson D., Rajadream M.A.,
RA      Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.,
RT      "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT      XIII."
RL      Nature 387:90-93 (1997).
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC      -1- SIMILARITY: TO YEAST YAL042W AND S.POMBE SPAC24B1.08C.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial/
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@lsb-sib.ch).
CC      -----
DR      EMBL, Z38114; CAAB6254.1; -.
DR      PIR, S48331; S48331.
DR      GenMonline, 142598; -.
DR      SGP, S0004532; BRV41.
DR      GO, GO:0030138; C:COPII-coated vesicle; IDA.
DR      GO, GO:0030176; C:integral to endoplasmic reticulum membrane; IDA.
DR      GO, GO:0030173; C:integral to Golgi membrane; IDA.
DR      GO, GO:0006888; P:ER to Golgi transport; IPI.

```


CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L00993; AAA39415.1; -
 DR EMBL; BC003820; AAH03820.1; -
 DR EMBL; Y07951; CAA69249.1; -
 DR MGD; MGI:98423; Ssb.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR InterPro; IPR002344; Lupus_La_dom.
 DR InterPro; IPR006630; RNA_rec_mot.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rtm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS0102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 111 187
 SQ SEQUENCE 415 AA; 47756 MW; 2D75197692FDC933 CRC64;
 Query Match 81.2%; Score 84.5; DB 1; Length 415;
 Best Local Similarity 94.4%; Pred. No. 2.1e-06;
 Matches 17; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 Oy 2 ALEAKICHOIEEYFGDF 19
 Db 12 ALEAKICHOI-EYFGDF 28
 RESULT 5
 LAB_XENLA STANDARD; PRT; 427 AA.
 ID LAB_XENLA STANDARD; PRT; 427 AA.
 AC P28049; 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein homolog B (La ribonucleoprotein B) (La autoantigen
 DE homolog B).
 GN LAB1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocyte;
 RX MEDLINE=93287095; PubMed=8510143;
 RA Scherly D., Stutz F., Lin-Marg N., Clarkson S.G.;
 RT "La proteins from Xenopus laevis. cDNA cloning and developmental
 RT expression.";
 RL J. Mol. Biol. 231:196-204(1993).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' terminus of virtually all nascent
 CC polymerase III transcripts (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DEVELOPMENTAL STAGE: Barely detectable in stage I/II oocytes,
 CC accumulate in stage III/IV oocytes, then exhibit a roughly
 CC constant steady state level in mature oocytes, eggs, and early
 CC embryos.
 CC -1- PTM: Phosphorylated (Probable).
 CC -1- MISCELLANEOUS: There are two forms of La, LaA and LaB, in xenopus.
 CC -1- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X68818; CAA8716.1; -
 DR PIR; S33817; S33817.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rtm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS0102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 110 202
 FT DOMAIN 315 331
 SQ SEQUENCE 427 AA; 48995 MW; 45F3146F934A355 CRC64;
 Query Match 57.2%; Score 59.5; DB 1; Length 427;
 Best Local Similarity 75.0%; Pred. No. 0.03;
 Matches 12; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
 Oy 3 LEAKICHOIEEYFGD 18
 Db 12 LDTKICEQI-EYFGD 26
 RESULT 6
 LAB_XENLA STANDARD; PRT; 428 AA.
 ID LAB_XENLA STANDARD; PRT; 428 AA.
 AC P28048; 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein homolog A (La ribonucleoprotein A) (La autoantigen
 DE homolog A).
 GN LAB1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocyte;
 RX MEDLINE=93287095; PubMed=8510143;
 RA Scherly D., Stutz F., Lin-Marg N., Clarkson S.G.;
 RT "La proteins from Xenopus laevis. cDNA cloning and developmental
 RT expression.";
 RL J. Mol. Biol. 231:196-204(1993).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' terminus of virtually all nascent
 CC polymerase III transcripts (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DEVELOPMENTAL STAGE: Barely detectable in stage I/II oocytes,
 CC accumulate in stage III/IV oocytes, then exhibit a roughly
 CC constant steady state level in mature oocytes, eggs, and early
 CC embryos.
 CC -1- PTM: Phosphorylated (Probable).
 CC -1- MISCELLANEOUS: There are two forms of La, LaA and LaB, in xenopus.
 CC -1- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

DR PROSITE: PS00030; RRM_RNP_1; 1.
 KW Systemic lupus erythematosus; RNA-binding; Phosphorylation;
 KM Nuclear protein.
 FT DOMAIN 111 187 RNA-BINDING (RRM).
 FT MOD_RES 366 366 PHOSPHORYLATION (BY CK2).
 SQ SEQUENCE 408 AA; 46837 MW; EC153C15F9187FC4 CRC64;

Query Match 85.1%; Score 88.5; DB 1; Length 408;
 Best Local Similarity 94.7%; Pred. No. 4.5e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 AALEAKICHQIEYYEGDF 19
 DB 11 AALEAKICHQI-EYYEGDF 28

RESULT 3

LA RAT ID LA RAT STANDARD; PRT; 415 AA.

AC P38556; 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
 homolog).
 GN SSB OR SS-B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN (1)

SEQUENCE FROM N.A.

WEDLINE=9324625; PubMed=7916708;
 RA Semel I., Trester H., Bartsch H., Schwemmler M., Igloi G.L.,
 RA Bachmann M.;
 RT "Isolation of rat cDNA clones coding for the autoantigen SS-B/La:
 RT detection of species-specific variations.";
 RL Gene 126:265-268(1993).

CC -1- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' terminus of virtually all nascent
 CC polymerase III transcripts. It is associated with precursor forms
 CC of RNA polymerase III transcripts including crRNA and 4.5S, 5S, 7S,
 CC and 7-2 RNAs.

CC -1- SUBUNIT: Interacts with DDX15 (By similarity).

CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).

CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X67859; CAA48043.1; -.

CC PIR; JCI494; JCI494.

CC InterPro; IPR002344; Lupus_La.

CC InterPro; IPR006630; Lupus_La_dom.

CC InterPro; IPR00504; RNA_rec_mot.

CC Pfam; PF00383; La; 1.

CC PRINTS; PR00302; LUPUSLA.

CC SMART; SMO0715; La; 1.

CC SMART; SMO0360; RRM; 1.

CC PROSITE; PS0102; RRM; 1.

CC PROSITE; PS00030; RRM_RNP_1; 1.

CC RNA-Binding; Nuclear protein; Phosphorylation.

CC DOMAIN 111 187 RNA-BINDING (RRM).
 SQ SEQUENCE 415 AA; 47777 MW; 033FD9CC18A75F98 CRC64;

Query Match 85.1%; Score 88.5; DB 1; Length 415;
 Best Local Similarity 94.7%; Pred. No. 4.6e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 AALEAKICHQIEYYEGDF 19
 DB 11 AALEAKICHQI-EYYEGDF 28

RESULT 4

LA MOUSE ID LA MOUSE STANDARD; PRT; 415 AA.

AC P32067; 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
 homolog).

GN SSB OR SS-B.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

NCBI_TaxID=10090;

FT DOMAIN 111 187 RNA-BINDING (RBM).
 SO SEQUENCE 404 AA; 46534 MW; 4EE30B5C262AD6A1 CRC64;
 Query Match 85.1%; Score 88.5; DB 1; Length 404;
 Best Local Similarity 94.7%; Pred. No. 4.5e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEEYFGDP 19
 |||||
 DB 11 AALEAKICHQI-EYFGDP 28

RESULT 2
 ID LA_HUMAN STANDARD; PRT; 408 AA.
 AC P05455;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein (Sjogren syndrome type B antigen) (SS-B) (La ribonucleoprotein) (La autoantigen).
 OS Homo sapiens (Human).
 GN Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OC NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=89202037; PubMed=2468131;
 RT Chan E.K.L., Sullivan K.F., Tan E.M.;
 RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences for RNA-binding.";
 RL Nucleic Acids Res. 17:2233-2244 (1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=89053970; PubMed=3192525;
 RT Chambers J.C., Kenan D., Martin B.J., Keene J.D.;
 RT "Genomic structure and amino acid sequence domains of the human La autoantigen.";
 RL J. Biol. Chem. 263:18043-18051 (1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta, and Skeletal muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stopleton M., Soares W.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Umed T.B., Toshiyuki S., Carinici P., Prange C., Rata S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullanb S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [4]
 RP SEQUENCE OF 54-408 FROM N.A.
 RX MEDLINE=8819081; PubMed=2452201;
 RA Sturgees A.D., Peterson M.G., McNeillage L.J., Whittingham S., Coppel R.S.;
 RT "Characteristics and epitope mapping of a cloned human autoantigen La.";
 RL J. Immunol. 140:3212-3218 (1988).
 RN [5]

RP SEQUENCE OF 54-97 FROM N.A.
 RX MEDLINE=8516283; PubMed=385688;
 RA Chambers J.C., Keene J.D.;
 RT "Isolation and analysis of cDNA clones expressing human lupus La antigen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2115-2119 (1985).
 RN [6]
 RP FUNCTION.
 RX MEDLINE=89251617; PubMed=2470590;
 RA Gottlieb E., Steltz J.A.;
 RT "Function of the mammalian La protein: evidence for its action in transcription termination by RNA polymerase III.";
 RL EMBO J. 8:851-861 (1989).
 RN [7]
 RP PHOSPHORYLATION.
 RX MEDLINE=97207017; PubMed=9054510;
 RA Fan H., Sakulich A.L., Goodier J.L., Zhang X., Qin J., Marate R.J.;
 RT "Phosphorylation of the human La antigen on serine 366 can regulate recycling of RNA polymerase III transcription complexes.";
 RL Cell 88:707-715 (1997).
 RN [8]
 RP INTERACTION WITH DDX15.
 RX MEDLINE=22346609; PubMed=12458796;
 RA Fouraux M.A., Kolman M.J.M., Van der Heijden A., De Jong A.S., Van Venrooij W.J., Pruijn G.J.M.;
 RT "The human La (SS-B) autoantigen interacts with DDX15/hPrp43, a putative DEAD-box RNA helicase.";
 RL RNA 8:1428-1443 (2002).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' termini of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S, and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE C-TERMINAL PART OF THE PROTEIN.
 CC -1- PTM: The N-terminus is blocked.
 CC -1- DISEASE: Sera from patients with systemic lupus erythematosus often contain antibodies that react with the normal cellular La protein as if this antigen was foreign.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RNM) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X13697; CAA31985.1; -;
 CC EMBL; J04205; AAA51885.1; -;
 CC EMBL; BC001289; AAH01289.1; -;
 CC EMBL; BC020818; AAH20818.1; -;
 CC PIR; A1888; A31888.
 CC GeneW; HGNC:11316; SSB.
 CC MIM; 109090; -;
 CC GO; GO:0030529; C:ribonucleoprotein complex; TAS.
 CC GO; GO:0003729; F:tRNA binding; TAS.
 CC GO; GO:0000049; F:tRNA binding; TAS.
 CC GO; GO:0008334; P:histone mRNA metabolism; TAS.
 CC GO; GO:0006400; P:tRNA modification; TAS.
 CC InterPro; IPR002344; Lupus La.
 CC InterPro; IPR006630; Lupus La dom.
 CC InterPro; IPR005054; RNA_rec_mot.
 CC Pfam; PF05383; La; 1.
 CC Pfam; PF00076; ttm; 1.
 CC PRINTS; PR00302; LUPUSLA.
 CC SMART; SM00715; LA; 1.
 CC SMART; SM00360; RRM; 1.
 CC PROSITE; PS50102; RRM; 1.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:01:41 ; Search time 7.64246 Seconds
(without alignments)
129.452 Million cell updates/sec

Title: US-09-836-073-16

```
Perfect score: 104
Sequence:      1 AALEAKICHQIEEYFGDF 19
```

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	88.5	85.1	404	1	LA_BOVIN	P10861 bos taurus
2	88.5	85.1	408	1	LA_HUMAN	P05455 homo sapien
3	88.5	85.1	415	1	LA_RAT	P38656 rattus norv
4	84.5	81.2	415	1	LA_MOUSE	P32067 mus musculu
5	59.5	57.2	427	1	LAB_XENIA	P28049 xenopus lae
6	55.5	53.4	428	1	LAB_XENIA	P28048 xenopus lae
7	47	45.2	482	1	LBP_FAB1T	P17454 oryctolagus
8	46	44.8	352	1	YMG7_YEAST	Q04651 saccharomyc
9	45.5	43.8	849	1	DOC6_MOUSE	Q8v4t9 mus musculu
10	45.5	43.8	2047	1	DOC6_HUMAN	Q96HP0 homo sapien
11	45	43.3	620	1	ACE1_CABR	Q27459 caenorhabdi
12	45	43.3	620	1	ACE1_CABEL	P38433 caenorhabdi
13	45	43.3	1164	1	RPO2_COMPLEX	P17474 comox vitu
14	45	43.3	1164	1	RPO2_VACCIN	P19798 vaccinia vir
15	45	43.3	1164	1	RPO2_VANU	P33811 variola vir
16	44.5	42.8	383	1	LA_AEDAL	Q26457 aedes albop
17	44	42.3	487	1	GATB_CHICV	P832W6 chlamydomo
18	44	42.3	552	1	GSTE_MYRPE	P35501 myzus persi
19	44	42.3	564	1	BSTF_MYRPE	P35502 myzus persi
20	43	41.3	88	1	Y419_NEIMA	Q91JDP5 neisseria m
21	43	41.3	297	1	MRT1_HUMAN	Q9H1U9 homo sapien
22	43	41.3	609	1	YAS8_MERJA	Q85458 methanococc
23	42	40.4	421	1	HMDH_AAREP	Q9Y8A6 aeropyrum p
24	42	40.4	428	1	Y813_TREPDA	Q83789 tireponema p
25	42	40.4	512	1	WATK_GINBI	Q8mx22 ginkgo bilo
26	42	40.4	1037	1	ACRD_ECOLI	P24177 escherichia
27	41	39.4	263	1	MTX2_MOUSE	Q88441 mus musculu
28	41	39.4	926	1	ME19_DROME	Q24087 drosophilila
29	40.5	38.9	390	1	LA_DROME	P40796 drosophilila
30	40.5	38.9	773	1	PNK4_HUMAN	Q9NUE7 homo sapien
31	40	38.5	263	1	MTX2_HUMAN	Q75431 homo sapien
32	40	38.5	271	1	YJ46_OCBITH	Q8EPY4 glycinebacil
33	40	38.5	320	1	RIAO_SOYBN	P50346 glycinib max

34	40	38.5	484	1	TH1_PSEBK	088c54 pseudomonas
33	40	38.5	509	1	STK_TYDAP	p17713 hydra attenu
32	40	38.5	571	1	PTI_CHLNP	Q92963 chlamydia p
31	40	38.5	597	1	MBHT_RHOQA	P32894 rhodobacter
30	40	38.5	620	1	SYR_BIFLO	Q85462 bifidobacter
29	40	38.5	812	1	MCW3_MOUSE	P25206 mus musculu
28	40	38.5	919	1	RPO2_CAPYK	P16716 capricipovir
27	40	38.5	919	1	POLS_SFV	P03315 semliki for
26	40	38.5	1253	1	AT7B_SHEEP	Q9X50 ovis aries
25	40	38.5	1505	1	PINK4_RAT	Q92308 ratu
24	39.5	38.0	773	1	DATA_STATN	P54654 staphylococ
23	39	37.5	282	1	HEM1_CHLCCV	Q82461 chlamydo
22	39	37.5	338	1		
21	39	37.5				
20	39	37.5				
19	39	37.5				
18	39	37.5				
17	39	37.5				
16	39	37.5				
15	39	37.5				
14	39	37.5				
13	39	37.5				
12	39	37.5				
11	39	37.5				
10	39	37.5				
9	39	37.5				
8	39	37.5				
7	39	37.5				
6	39	37.5				
5	39	37.5				
4	39	37.5				
3	39	37.5				
2	39	37.5				
1	39	37.5				

ALIGNMENTS

RESULT 1	LA BOVIN	STANDARD;	PRT;	404 AA.
ID_	LA BOVIN			
AC	P10881;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-JUL-1989 (Rel. 11, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Lupus La protein homolog (La ribonucleoprotein) (La autoantigen homolog).			
DE	SSB.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euteheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Plutitary;			
RX	MEDLINE=69202037; PubMed=2468131;			
RA	Chan E.K.U., Sullivan K.F., Tan E.M.;			
RT	"Ribonucleoprotein SS-/La belongs to a protein family with consensus sequences for RNA-binding.";			
RT	Nucleic Acids Res. 17:2233-2244(1989).			
RL	-1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' termini of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including crRNA and 4.5S, 5S, 7S, and 7-2 RNAs.			
CC	-1- SUBUNIT: Interacts with DDX15 (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Nuclear (Probable).			
CC	-1- PM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE C-TERMINAL PART OF THE PROTEIN.			
CC	-1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.			
CC	-----			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-bdb.ch/announce/ or send an email to license@isb-bdb.ch).			
CC	-----			
DR	EMBL; X13698; CAA31986.1; -.			
DR	PIR; S03849; S03849.			
DR	InterPro; IPR002344; lupus_La.			
DR	InterPro; IPR006630; lupus_La_dom.			
DR	InterPro; IPR000504; RNA_rec_mot.			
DR	Pfam; PF05383; La; 1.			
DR	Pfam; PF00076; rrm; 1.			
DR	PRINTS; PR00302; LUPUSLA.			
DR	SMART; SM00715; LA; 1.			
DR	SMART; SM00360; RRM; 1.			
DR	PROSITE; PS50102; RRM; 1.			
DR	PROSITE; PS00030; RRM_RNP_1; 1.			
KW	RNA-binding; Nuclear protein; Phosphorylation.			

This Page Blank (uspte)

A:Reference number: Z20693
 A:Accession: T29824
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-620 <WUX>
 A:Cross-references: EMBL:U58731; PIDN:AAB00593.1; GSPDB:GN00028
 A:Experimental source: strain Bristol N2; clone W09B12
 C:Genetics:
 A:Gene: CESP:ace-1
 A:Map position: X
 A:Introns: 13/3; 59/2; 154/2; 236/3; 296/2; 454/2; 509/2; 573/1; 606/2
 C:Superfamily: cholinesterase; cholinesterase homology
 C:Keywords: carboxylic ester hydrolase
 F:45-567/Domain: cholinesterase homology <CHE>

Query Match 43.3%; Score 45; DB 2; Length 620;
 Best Local Similarity 66.7%; Pred. No. 37;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 8 CHOIEHYFGDF 19
 | | | | |
 Db 82 CIQSEDTYFGDF 93

RESULT 15
 T37411
 RNA polymerase subunit rpo132 - vaccinia virus (strain Ankara)
 C:Species: vaccinia virus
 A:Variety: strain Ankara
 C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000
 C:Accession: T37411
 R:Antoine, G.; Schefflinger, F.; Falkner, F.G.; Dörner, F.
 submitted to the EMBL Data Library, March 1997
 A:Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain
 A:Reference number: Z20877
 A:Accession: T37411
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1156 <ANT>
 A:Cross-references: EMBL:U94848; PIDN:AAB96526.1
 A:Experimental source: strain Ankara
 C:Genetics:
 A:Note: MVA135R
 C:Superfamily: DNA-directed RNA polymerase 132k polypeptide

Query Match 43.3%; Score 45; DB 2; Length 1156;
 Best Local Similarity 56.2%; Pred. No. 66;
 Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 LEAKICHOIEHYFGD 18
 | | | | |
 Db 497 LEKKICEYIRSYKDD 512

Search completed: September 10, 2004, 18:02:35
 Job time : 16.8603 secs

lipopolysaccharide-binding protein - rabbit
 C:Species: *Oryctolagus cuniculus* (domestic rabbit)
 C:Date: 05-Oct-1990 #sequence_revision 05-Oct-1990 #text_change 20-Aug-1999
 C:Accession: B35843; A46553
 R:Schumann, R.R.; Leong, S.R.; Flaegge, G.W.; Gray, P.W.; Wright, S.D.; Mathison, J.C.; J
 Science 249, 1429-1431, 1990
 A>Title: Structure and function of lipopolysaccharide binding protein.
 A:Reference number: A35843; MUID:90385281; PMID:2402637
 A:Accession: B35843
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-482 <SC>
 A:Cross-references: GB:M35534; NID:G165467; PIDN:AAA99235.1; PID:G165468
 R:Tobias, P.S.; Soldau, K.; Ulevitch, R.J.
 J. Exp. Med. 164, 777-793, 1986
 A>Title: Isolation of a lipopolysaccharide-binding acute phase reactant from rabbit seru
 A:Reference number: A46553; MUID:86306528; PMID:2427635
 A:Accession: A46553
 A:Molecule type: protein
 A:Residues: 27-55, 'XG', 58-62, 'F', 64-65 <TOB>
 C:Superfamily: lipopolysaccharide-binding protein
 C:Keywords: acute phase

Query Match 45.2%; Score 47; DB 2; Length 482;
 Best Local Similarity 81.8%; Pred. No. 14;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 LEAKICHOIEB 13
 ||:|||||
 Db 194 LESKICRQIEE 204

RESULT 11
 S48331
 probable membrane protein YML067C - yeast (*Saccharomyces cerevisiae*)
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 19-Apr-2002
 C:Accession: S48331; S48330
 R:Bowman, S.; Churcher, C.
 submitted to the EMBL Data Library, August 1994
 A:Reference number: S48326
 A:Accession: S48331
 A:Molecule type: DNA
 A:Residues: 1-352 <BOW>
 A:Cross-references: EMBL:Z38114; NID:G558402; PID:G558408; GSPDB:GN00013; MIPS:YML067C
 A:Accession: S48330
 A:Molecule type: DNA
 A:Residues: 69-352 <BOF>
 A:Cross-references: EMBL:Z38114; NID:G558402; PID:G558407
 C:Genetics:
 A:Gene: SGD:ERV41; MIPS:YML067C
 A:Cross-references: SGD:S0004532
 A:Map position: 13L
 A:Intron: 11/1
 C:Keywords: transmembrane protein
 F:230-246/Domain: transmembrane #status predicted <TM1>
 F:248-264/Domain: transmembrane #status predicted <TM2>

Query Match 44.2%; Score 46; DB 2; Length 352;
 Best Local Similarity 56.2%; Pred. No. 15;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 4 EAKICHOIEBYFGDF 19
 ||:|||||
 Db 196 ELKFNHVINERFSFGDF 211

RESULT 12
 B69391
 hypothetical protein AF1134 - *Archaeoglobus fulgidus*
 C:Species: *Archaeoglobus fulgidus*
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: B69391

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.;
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A>Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: B69391
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-456 <KLE>
 A:Cross-references: GB:AE001026; GB:AE000782; NID:G2689349; PIDN:AB90123.1; PID:G264946;

Query Match 44.2%; Score 46; DB 2; Length 456;
 Best Local Similarity 61.5%; Pred. No. 19;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 LEAKICHOIEBY 15
 ||:|||||
 Db 405 WEAKIPHOIEBY 417

RESULT 13
 H90096
 eukaryotic release factor 1 homolog [imported] - *Gulliardia theta* nucleomorph
 C:Species: nucleomorph *Gulliardia theta*
 A>Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 15-Jun-2001
 C:Accession: H90096
 R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Rel
 Nature 410, 1091-1096, 2001
 A>Title: The highly reduced genome of an enslaved algal nucleus.
 A:Reference number: A99082; MUID:11323671; PMID:11323671
 A:Accession: H90096
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-409 <DOU>
 A:Cross-references: GB:AF165818; NID:G13794528; PIDN:AAK39903.1; GSPDB:GN00150
 C:Genetics:
 A:Gene: erf1
 A:Map position: 1
 A:Genome: nucleomorph
 C:Superfamily: translation releasing factor erf-1
 C:Keywords: nucleomorph

Query Match 43.3%; Score 45; DB 2; Length 409;
 Best Local Similarity 50.0%; Pred. No. 25;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 6 KICHOIEBYFGDF 19
 ||:|||||
 Db 200 KICEADQYLSDF 213

RESULT 14
 A54413
 acetylcholinesterase (EC 3.1.1.7) A precursor - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 21-Jul-2000
 C:Accession: A54413; T29824
 R:Arpagaus, M.; Fedon, Y.; Cousin, X.; Chatonnet, A.; Berge, J.B.; Fournier, D.; Toutant
 J. Biol. Chem. 269, 9957-9965, 1994
 A>Title: cDNA sequence, gene structure, and in vitro expression of ace-1, the gene encod
 A:Reference number: A54413; MUID:94193691; PMID:8144590
 A:Accession: A54413
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-620 <ARP>
 A:Cross-references: GB:X75311; NID:G475060; PIDN:CAA3080.1; PID:G671831
 R:Wu, X.; Le, T.T.
 submitted to the EMBL Data Library, May 1996
 A:Description: The sequence of C. elegans cosmid W09B12.

R.Scherly, D.; Stutz, F.; Lin-Marg, N.; Clarkson, S.G.

J. Mol. Biol. 231, 196-204, 1993

A>Title: La proteins from *Xenopus laevis*. cDNA cloning and developmental expression.

A/Reference number: S33817; MUID:93287095; PMID:8510143

A/Accession: S33818

A/Molecule type: mRNA

A/Residues: 1-428 <SCH>

A/Cross-references: EMBL:X68817; NID:g64873; PID:CAA48715.1; PID:g64874

C/Comment: This protein associates with a variety of small RNA molecules, most of which

act as a transcription termination factor.

C/Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology

C/Keywords: phosphoprotein; RNA binding

F.112-178/Domain: ribonucleoprotein repeat homology <RRM>

F.113-118/Region: RNA-binding RNP motif

F.151-158/Region: RNA-binding RNP motif

F.228-428/Domain: phosphorylated #status predicted <PHY>

Query Match 53.4%; Score 55.5; DB 1; Length 428;
Best Local Similarity 73.3%; Pred. No. 0.57;
Matches 11; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

RESULT 6
B70456
thioredoxin - Aquifex aeolicus

C/Species: Aquifex aeolicus

C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999

C/Accession: B70456

R/Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Oy

V.

Nature 392, 353-358, 1998

A>Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.

A/Reference number: A70300; MUID:9819666; PMID:9537320

A/Accession: B70456

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-135 <NOF>

A/Cross-references: GB:AE000757; NID:g2984092; PID:AA07635.1; PID:g2984097; GB:AE00065

A/Experimental source: strain VFS

C/Genetics:

Query Match 52.9%; Score 55; DB 2; Length 135;
Best Local Similarity 72.7%; Pred. No. 0.23;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 CHQIEEYFGD 18
|||:|:|:|:|
Db 52 CHQIEEYFGD 62

RESULT 7
AC1870
hypothetical protein alr0508 [imported] - Nostoc sp. (strain PCC 7120)

C/Species: Nostoc sp. PCC 7120

A/Note: Nostoc sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C/Accession: AC1870

R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi,

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 9, 205-213, 2001

A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ana*

A/Reference number: AB1807; MUID:21595285; PMID:11759840

A/Accession: AC1870

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-187 <KUR>
A/Cross-references: GB:BA000019; PID:BA072466.1; PID:gl7129853; GSPDB:GN00179
A/Experimental source: strain PCC 7120

C/Genetics:
A/Gene: alr0508

Query Match 45.2%; Score 47; DB 2; Length 187;
Best Local Similarity 50.0%; Pred. No. 5.6;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHQIEEYFGD 18
|||:|:|:|:|
Db 97 ICHQIEEYFGD 108

RESULT 8
H90120
hypothetical protein orf357 [imported] - *Guillardia theta* nucleomorph

C/Species: nucleomorph *Guillardia theta*

A/Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont

C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001

C/Accession: H90120

R/Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Re

Nature 410, 1091-1096, 2001

A>Title: The highly reduced genome of an enslaved algal nucleus.

A/Reference number: A99082; MUID:11323671; PMID:11323671

A/Accession: H90120

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-357 <DOU>

A/Cross-references: GB:AF083031; NID:gl3794284; PID:AAK39661.1; GSPDB:GN00152

C/Genetics:

A/Gene: orf357

A/Map position: 3

A/Genome: nucleomorph

C/Keywords: nucleomorph

Query Match 45.2%; Score 47; DB 2; Length 357;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 LEAKICHOIEEYFGD 18
|:|:|:|:|:|:|
Db 242 LKNKVFQNMKXYFGD 257

RESULT 9
T39314
hypothetical protein SPBC119.16c - fission yeast (*Schizosaccharomyces pombe*)

C/Species: *Schizosaccharomyces pombe*

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C/Accession: T39314

R/Mood, V.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.

submitted to the EMBL Data Library, March 1998

A/Reference number: Z21843

A/Status: preliminary; translated from GB/EMBL/DDBJ

A/Molecule type: DNA

A/Residues: 1-448 <MOO>

A/Cross-references: EMBL:AL022117; PID:CAA17931.1; GSPDB:GN00067; SPDB:SPBC119.16c

A/Experimental source: strain 972n-; cosmid c119

C/Genetics:

A/Gene: SPDB:SPBC119.16c

Query Match 45.2%; Score 47; DB 2; Length 448;
Best Local Similarity 47.4%; Pred. No. 13;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 ALEAKICHOIEEYFGD 19
|:|:|:|:|:|:|
Db 9 ATPDAKEGVNENYFGD 27

RESULT 10
B35843

A:Reference number: S03848; MUID:89202037; PMID:2468131
 A:Accession: S03848
 A:Molecule type: mRNA
 A:Residues: 1-408 <CH2>
 A:Cross-references: EMBL:X13697; NID:g36414; PIDN:CAA31985.1; PID:g36415
 R:Chambers, J.C.; Keene, J.D.
 Proc. Natl. Acad. Sci. U.S.A. 82, 2115-2119, 1985
 A>Title: Isolation and analysis of cDNA clones expressing human lupus La antigen.
 A:Reference number: A22956; MUID:85166283; PMID:3856888
 A:Accession: A22956
 A:Molecule type: mRNA
 A:Residues: 45-97, 'LK' <CH3>
 A:Cross-references: GB:J06205
 A>Note: This sequence has been revised in reference A31888
 R:Myman, U.; Ringertz, N.R.; Petersson, I.
 Immunol. Lett. 22, 65-72, 1989
 A>Title: Demonstration of an amino terminal La epitope recognized by human anti-La sera.
 A:Reference number: A61051; MUID:89379261; PMID:2476379
 A:Accession: A61051
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-19, 'E', '21-47' <NYM>
 R:Sturges, A.D.; Peterson, M.G.; McNeillage, L.J.; Whittingham, S.; Coppel, R.L.
 J. Immunol. 140, 3212-3218, 1988
 A>Title: Characteristics and epitope mapping of a cloned human autoantigen La.
 A:Reference number: S11013; MUID:88199081; PMID:2452201
 A:Accession: S11013
 A:Molecule type: mRNA
 A:Residues: 'E', '55-287', 'V', '289-408' <STU>
 A:Cross-references: EMBL:M20328; NID:g337456; PIDN:AAA36577.1; PID:g337457
 R:Kobayashi, H.; Yamamoto, K.; Fujii, H.; Mura, H.; Miyasaka, K.; Miyamoto, K.
 J. Clin. Invest. 85, 1566-1574, 1990
 A>Title: Fine epitope mapping the human SS-B/La protein. Identification of a distinct
 A:Reference number: I55553; MUID:90237237; PMID:1692037
 A:Accession: I55553
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 81-107 <RES>
 A:Cross-references: GB:M35261; NID:g338491; PIDN:AAA36652.1; PID:g338495
 A:Accession: I70205
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 174-224 <RE2>
 A:Cross-references: GB:M35263; NID:g338492; PIDN:AAA36653.1; PID:g338496
 A:Accession: I70206
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 279-342 <RE3>
 A:Cross-references: GB:M35262; NID:g338493; PIDN:AAA36654.1; PID:g338497
 C:Comment: This protein associates with a variety of small RNA molecules, most of which
 act as a transcription termination factor.
 C:Keywords: phosphoprotein; RNA binding
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 F:112-178/Domain: ribonucleoprotein repeat homology <RPM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:228-408/Domain: phosphotyrlated #status experimental <PHY>

Query Match 85.1%; Score 88.5; DB 1; Length 408;
 Best Local Similarity 94.7%; Pred. No. 3.5e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHOIEYYFGDF 19
 |||||
 Db 11 AALEAKICHOI-EYYFGDF 28

RESULT 3

JC1494
 ribonucleoprotein La - rat
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 24-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: JC1494; S25145
 R:Semel, I.; Troster, H.; Bartsch, H.; Schwemmler, M.; Igloi, G.L.; Bachmann, M.
 Gene 126, 265-268, 1993
 A>Title: Isolation of rat cDNA clones coding for the autoantigen SS-B/La: Detection of
 A:Reference number: JC1494; MUID:93246255; PMID:7916708
 A:Accession: JC1494
 A:Molecule type: mRNA
 A:Residues: 1-415 <SEM>
 A:Cross-references: GB:X67859; NID:g55778; PIDN:CAA48043.1; PID:g55779
 A:Experimental source: liver
 C:Comment: This protein associates with a variety of small RNA molecules, most of which
 act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:112-178/Domain: ribonucleoprotein repeat homology <RPM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:227-415/Domain: phosphotyrlated #status predicted <PHY>

Query Match 85.1%; Score 88.5; DB 1; Length 415;
 Best Local Similarity 94.7%; Pred. No. 3.6e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHOIEYYFGDF 19
 |||||
 Db 11 AALEAKICHOI-EYYFGDF 28

RESULT 4

S33817
 ribonucleoprotein La-B - African clawed frog
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33817; S28544
 R:Schery, D.; Stutz, F.; Lin-Marg, N.; Clarkson, S.G.
 J. Mol. Biol. 231, 196-204, 1993
 A>Title: La proteins from Xenopus laevis. cDNA cloning and developmental expression.
 A:Reference number: S33817; MUID:93287095; PMID:8510143
 A:Accession: S33817
 A:Molecule type: mRNA
 A:Residues: 1-427 <SCH>
 A:Cross-references: EMBL:X68818; NID:g64875; PIDN:CAA48716.1; PID:g64876
 C:Comment: This protein associates with a variety of small RNA molecules, most of which
 act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:111-177/Domain: ribonucleoprotein repeat homology <RPM>
 F:112-117/Region: ribonucleoprotein repeat homology
 F:150-157/Region: RNA-binding RNP2 motif
 F:227-427/Domain: phosphotyrlated #status predicted <PHY>

Query Match 57.2%; Score 59.5; DB 1; Length 427;
 Best Local Similarity 75.0%; Pred. No. 0.13;
 Matches 12; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 3 LEAKICHOIEYYFGDF 18
 |||||
 Db 12 LDTKICEOI-EYYFGDF 26

RESULT 5

S33818
 ribonucleoprotein La-A - African clawed frog
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33818; S28545

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:33:11; Search time 14.8603 Seconds
(without alignments)
122.988 Million cell updates/sec

Title: US-09-836-073-16

Perfect score: 104

Sequence: 1 AALEAKICHIQEYFGDF 19

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

1: PIR 78: *
2: p1r1: *
3: p1r2: *
4: p1r3: *
5: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88.5	85.1	404	1 S03849	ribonucleoprotein
2	88.5	85.1	408	1 A31888	ribonucleoprotein
3	88.5	85.1	415	1 J01494	ribonucleoprotein
4	59.5	57.2	427	1 S33817	ribonucleoprotein
5	55.5	53.4	428	1 S33818	ribonucleoprotein
6	55	52.9	135	2 B70456	thioredoxin - Aqu
7	47	45.2	187	2 AC1870	hypothetical prote
8	47	45.2	357	2 H90120	hypothetical prote
9	47	45.2	448	2 T39314	hypothetical prote
10	47	45.2	482	2 B35843	lipopolysaccharide
11	46	44.2	352	2 S48331	probable membrane
12	46	44.2	456	2 E69391	hypothetical prote
13	45	43.3	409	2 H90086	eukaryotic release
14	45	43.3	620	2 A54413	acetylcholinestera
15	45	43.3	1156	2 T37411	RNA polymerase sub
16	45	43.3	1164	1 RNW28T	DNA-directed RNA p
17	45	43.3	1164	1 RNW28T	DNA-directed RNA p
18	45	43.3	1164	2 T28566	DNA-directed RNA p
19	45	43.3	1164	2 F72166	A25R protein - var
20	45	43.3	1164	2 G36850	A24R protein - var
21	44	42.3	552	2 S36786	carboxylesterase (
22	44	42.3	564	2 S36787	carboxylesterase (
23	43	41.3	88	2 H81014	conserved hypotet
24	43	41.3	427	2 G70590	probable dead3 pro
25	43	41.3	510	2 T06495	glucose-1-phosphat
26	43	41.3	609	2 A64432	modulator factor
27	42	40.4	386	2 H86870	hypothetical prote
28	42	40.4	421	2 E72573	probable 3-hydroxy
29	42	40.4	428	2 B71278	hypothetical prote

30	42	40.4	468	2 T21535	hypothetical prote
31	42	40.4	541	2 H71887	hypothetical prote
32	42	40.4	542	2 G64627	hypothetical prote
33	42	40.4	905	2 T38314	probable vacuolar
34	42	40.4	969	2 A75634	McR8-related prote
35	42	40.4	1037	2 E65022	acridinyl resistin
36	42	40.4	1037	2 D91045	aminoglycoside eff
37	42	40.4	1037	2 G85889	hypothetical prote
38	42	40.4	1069	2 AF1930	hypothetical prote
39	41.5	39.9	463	2 T39004	probable histone a
40	41	39.4	153	2 T19054	hypothetical prote
41	41	39.4	203	2 T28352	ORF MSV191 MFG mot
42	41	39.4	528	2 T02972	SRP protein homol
43	41	39.4	650	2 C69678	involved in polyke
44	41	39.4	771	2 T43612	transposase - Yers
45	41	39.4	926	2 S58936	meiotic recombinat

ALIGNMENTS

RESULT 1
S03849
ribonucleoprotein Ia - bovine
N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
C:Species: Bos primigenius taurus (cattle)
C>Date: 31-Dec-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C:Accession: S03849
R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.
Nucleic Acids Res. 17, 2233-2244, 1989
A>Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences
A:Reference number: S03848; MWID:89202037; PMID:2468131
A:Accession: S03849
A:Molecule type: mRNA
A:Residues: 1-404 <CHA>
A:Cross-references: EMBL:X13698; NID:g755; PIDN:CAA31986.1; PID:g756
A:Note: part of this sequence was confirmed by protein sequencing
A:Comment: This protein associates with a variety of small RNA molecules, most of which
act as a transcription termination factor.
C:Superfamily: ribonucleoprotein Ia; ribonucleoprotein repeat homology
C:Keywords: blocked amino end; phosphoprotein; RNA binding
F:112-178/Domain: ribonucleoprotein repeat homology <RKM>
F:113-118/Region: RNA-binding RNP2 motif
F:151-158/Region: RNA-binding RNP2 motif
F:228-404/Domain: phosphorylated #statue predicted <PHY>

Query Match
Best Local Similarity 85.1%; Score 88.5; DB 1; Length 404;
Matches 18; Conservativity 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 AALEAKICHIQEYFGDF 19
Db 11 AALEAKICHIQEYFGDF 28

RESULT 2
A31888
ribonucleoprotein Ia - human
N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B; Sjogren syndrome antigen
C:Species: Homo sapiens (man)
C>Date: 21-May-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C:Accession: A31888; S03848; A22956; A61051; S11013; I55553; I70205; I70206; A31273
R:Chambers, J.C.; Kenan, D.; Martin, B.J.; Keene, J.D.
J. Biol. Chem. 263, 18043-18051, 1988
A>Title: Genomic structure and amino acid sequence domains of the human Ia autoantigen.
A:Reference number: A31888; MWID:89053970; PMID:3192525
A:Accession: A31888
A:Molecule type: mRNA
A:Residues: 1-408 <CHA>
A:Cross-references: GB:J04205; NID:g178686; PIDN:AAA51885.1; PID:g178687
R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.
Nucleic Acids Res. 17, 2233-2244, 1989
A>Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences

This Page Blank (uspto)

Mon Sep 13 09:36:13 2004

us-09-836-073-15.rapb

Search completed: September 10, 2004, 18:11:56
Job time : 41.424 secs

```

: FILE REFERENCE: 22000205482
: CURRENT APPLICATION NUMBER: US/09/836,073
: PRIOR FILING DATE: 2002-10-24
: PRIOR APPLICATION NUMBER: 09/316,630
: PRIOR FILING DATE: 1999-05-21
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 9
: LENGTH: 18
: TYPE: prt
: ORGANISM: Homo Sapiens
: US-09-836-073-9

```

Query Match	70.0%	Score 70;	DB 9;	Length 18
Best Local Similarity	80.0%	Pred. No. 0.00075;		
Matches 12; Conservative		1; Mismatches	2; Indels	

OY 3 LDTKICEQIEYYFGD 17
 | : |||||
Db 3 LEAKICHOIEYYFGD 17

```

RESULT 12 073-11
US-09-836-073-11
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Dasg, S.
; APPLICANT: Baluya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-09-836-073-11

```

```

Query Match      68.0%; Score 68; DB 9; Length 18;
Best Local Similarity 75.0%; Pred. No. 0.0015;
Matches 12; Conservative 1; Mismatches 3; Indels 0;
QY      3 LDTKICGEQIERYFGDF 18
         : ||| ||| ||| |||
Db       3 LEAKICHQIEQYFGDF 18

```

```

RESULT 13
US-09-836-073-12
Sequence 12, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Das, S.
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 22002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ. ID NOS.: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 18
TYPE: PRT
ORGANISM: Homo Sapiens

```

US-09-836-073-12

Query Match	68.0%;	Score 68;	DB 9;	Length 18;
Best Local Similarity	75.0%;	Pred. No. 0.0015;		
Matches 12;	Conservative 1;	Mismatches 3;	Indels 0;	Gaps 0;

```
QY      3 LDTRICEQIEYYFGDF 18
        | : ||| ||| |||
Db      3 LEAKICHQIEYQFGDF 18
```

```

RESULT 14
US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baluya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: prt
; ORGANISM: Homo Sapiens
US-09-836-073-10

```

Query Match	67.0%	Score 67;	DB 9;	Length 18;
Best Local Similarity	75.0%	Pred. No. 0.0022;		
Matches 12;	Conservative 1;	Mismatches 3;	Indels 0;	Gaps 0;

```
QY      3 LDTKICEQIEYFGEF 18
      | : ||| ||| |||
Db      3 LEAKICHQIEYYQGDF 18
```

```

RESULT 15-073-16
US-09-836-073-16
? Sequence 16 Application US/09836073
? Patent No. US20020173475A1
? GENERAL INFORMATION:
? APPLICANT: Dasgupta, Asim
? APPLICANT: Das, S.
? APPLICANT: Baldya, Narayan
? TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
? FILE REFERENCE: 220002054822
? CURRENT APPLICATION NUMBER: US/09/836,073
? CURRENT FILING DATE: 2002-10-24
? PRIOR APPLICATION NUMBER: 09/316,630
? PRIOR FILING DATE: 1999-05-21
? NUMBER OF SEQ ID NOS: 19
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 16
? LENGTH: 19
? TYPE: prt
? ORGANISM: Rat
US-09-836-073-16

```

Query Match	65.5%	Score 65.5	DB 9.5	Length 19
Best Local Similarity	76.5%	Pred. No. 0.0039		
Matches 13; Conservative	1	Mismatches 2	Indels 1	Gaps 1
QY	3	LDTKICEQI-EYRGDF	18	
	:			
Db	3	LEATKIQIEEYRGDF	19	

APPLICANT: Das, S.
APPLICANT: Baidya, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

```
APPLICANT: Baluya, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 220002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13
LENGTH: 17
TYPE: PRT
ORGANISM: Mouse
US-09-836-073-13
```

```
Query Match          76.0%; Score 76; DB 9; Length 17;
Best Local Similarity 81.2%; Pred. No. 8,6e-05;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      3 LDTKICQIEYFGDF 18
        |:|||||
Db      2 LEAKICQIEYFGDF 17
```

```
RESULT 3
US-09-836-073-1
Sequence 1, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Das, S.
APPLICANT: Baluya, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 220002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 18
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-836-073-1
```

```
Query Match          76.0%; Score 76; DB 9; Length 18;
Best Local Similarity 81.2%; Pred. No. 9,1e-05;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      3 LDTKICQIEYFGDF 18
        |:|||||
Db      3 LEAKICQIEYFGDF 18
```

```
RESULT 4
US-09-836-073-14
Sequence 14, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Das, S.
APPLICANT: Baluya, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 220002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 18
```

```
TYPE: PRT
ORGANISM: Bovine
US-09-836-073-14
```

```
Query Match          76.0%; Score 76; DB 9; Length 18;
Best Local Similarity 81.2%; Pred. No. 9,1e-05;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      3 LDTKICQIEYFGDF 18
        |:|||||
Db      3 LEAKICQIEYFGDF 18
```

```
RESULT 5
US-10-170-385-477
Sequence 477, Application US/10170385
Publication No. US20030203372A1
GENERAL INFORMATION:
APPLICANT: Ward, Neil Raymond
APPLICANT: Mundy, Christopher Robert
APPLICANT: Kan, On
APPLICANT: Harris, Robert Alan
APPLICANT: White, Jonathan
APPLICANT: Binley, Katie Mary
APPLICANT: Rayner, William Nigel
APPLICANT: Naylor, Stuart
APPLICANT: Kingsman, Susan Mary
APPLICANT: Krige, David
TITLE OF INVENTION: ANALYSIS METHOD
FILE REFERENCE: 532682000100
CURRENT APPLICATION NUMBER: US/10/170,385
CURRENT FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: PCT/GB02/01662
PRIOR FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: PCT/GB01/05458
PRIOR FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 549
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 477
LENGTH: 408
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-170-385-477
```

```
Query Match          76.0%; Score 76; DB 12; Length 408;
Best Local Similarity 81.2%; Pred. No. 0.0023;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      3 LDTKICQIEYFGDF 18
        |:|||||
Db      13 LEAKICQIEYFGDF 28
```

```
RESULT 6
US-09-925-298-695
Sequence 695, Application US/09925298
Publication No. US20020039764A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA103
CURRENT APPLICATION NUMBER: US/09/925,298
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05881
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 846
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 695
LENGTH: 460
TYPE: PRT
ORGANISM: Homo sapiens
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:52:06 ; Search time 40.324 Seconds
(without alignments)
143.151 Million cell updates/sec

Title: US-09-836-073-15

Perfect score: 100
Sequence: 1 LDDTRKICEQIERYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/prodata/1/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubppaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/1/pubppaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/1/pubppaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubppaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/1/pubppaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/1/pubppaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/1/pubppaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep.*
10: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/1/pubppaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/prodata/1/pubppaa/US09C_NEW_PUB.pep.*
13: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep.*
14: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep.*
15: /cgn2_6/prodata/1/pubppaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata/1/pubppaa/US10C_NEW_PUB.pep.*
17: /cgn2_6/prodata/1/pubppaa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/1/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	100.0	18	US-09-836-073-15	Sequence 15, Appl
2	76	76.0	17	US-09-836-073-13	Sequence 13, Appl
3	76	76.0	18	US-09-836-073-1	Sequence 1, Appl
4	76	76.0	18	US-09-836-073-14	Sequence 14, Appl
5	76	76.0	408	US-10-170-385-477	Sequence 477, App
6	76	76.0	460	US-09-925-298-695	Sequence 695, App
7	76	76.0	460	US-10-102-806-695	Sequence 695, App
8	76	76.0	460	US-10-264-049-2643	Sequence 2643, App
9	74	74.0	18	US-09-836-073-2	Sequence 2, Appl
10	72	72.0	18	US-09-836-073-4	Sequence 4, Appl
11	70	70.0	18	US-09-836-073-9	Sequence 9, Appl
12	68	68.0	18	US-09-836-073-11	Sequence 11, Appl
13	68	68.0	18	US-09-836-073-12	Sequence 12, Appl
14	67	67.0	18	US-09-836-073-10	Sequence 10, Appl
15	65.5	65.5	19	US-09-836-073-16	Sequence 16, Appl

16	65	65.0	18	US-09-836-073-3	Sequence 3, Appl
17	65	65.0	21	US-10-376-121A-20	Sequence 20, Appl
18	64	64.0	18	US-09-836-073-5	Sequence 5, Appl
19	64	64.0	38	US-09-843-676-25	Sequence 25, Appl
20	64	64.0	38	US-09-766-253-25	Sequence 25, Appl
21	64	64.0	38	US-09-438-466-25	Sequence 25, Appl
22	64	64.0	38	US-10-325-810-215	Sequence 215, App
23	64	64.0	38	US-10-053-758-25	Sequence 25, Appl
24	64	64.0	38	US-10-054-295-25	Sequence 25, Appl
25	64	64.0	38	US-10-054-611-25	Sequence 7, Appl
26	61	61.0	18	US-09-836-073-7	Sequence 8, Appl
27	60	60.0	18	US-09-836-073-8	Sequence 214, App
28	60	60.0	38	US-10-325-810-214	Sequence 254661, Sequence 254664, Sequence 17, Appl
29	56	56.0	143	US-10-424-599-254661	Sequence 19, Appl
30	56	56.0	395	US-10-424-599-254664	Sequence 8, Appl
31	55	55.0	18	US-09-836-073-17	Sequence 45759, A
32	54	54.0	16	US-09-836-073-8	Sequence 71434, A
33	53	53.0	14	US-10-177-478-8	Sequence 24, Appl
34	52	52.0	303	US-10-425-114-45759	Sequence 24, Appl
35	52	52.0	401	US-10-425-114-45759	Sequence 24, Appl
36	51	51.0	377	US-10-424-599-254663	Sequence 24, Appl
37	51	51.0	488	US-10-424-599-272690	Sequence 125036, Sequence 49433, A
38	50.5	50.5	37	US-09-843-676-24	
39	50.5	50.5	37	US-09-766-253-24	
40	50.5	50.5	37	US-09-438-466-24	
41	50.5	50.5	37	US-10-054-758-24	
42	50.5	50.5	37	US-10-054-295-24	
43	50.5	50.5	37	US-10-054-611-24	
44	50	50.0	922	US-10-437-963-125036	
45	49	49.0	329	US-10-425-114-49433	

ALIGNMENTS

RESULT 1
US-09-836-073-15
; Sequence 15, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; TYPE: PRT
; LENGTH: 18
; ORGANISM: Xenopus
US-09-836-073-15

Query Match 100.0%; Score 100; DB 9; Length 18;

Best Local Similarity 100.0%; Pred. No. 2e-08; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LDDTRKICEQIERYFGDF 18
Db 1 LDDTRKICEQIERYFGDF 18

RESULT 2
US-09-836-073-13
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-24

Query Match 50.5%; Score 50.5; DB 3; Length 37;
Best Local Similarity 83.3%; Pred. No. 0.21;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Oy 7 ICEQIEYFGDF 18
||| |||||
Db 1 ICHQ-EYFGDF 11

RESULT 15
US-09-430-323-24
Sequence 24, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin
Andrews, William H.
TITLE OF INVENTION: No. 6309867e1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-430-323-24

Query Match 50.5%; Score 50.5; DB 4; Length 37;
Best Local Similarity 83.3%; Pred. No. 0.21;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Oy 7 ICEQIEYFGDF 18
||| |||||
Db 1 ICHQ-EYFGDF 11

Search completed: September 10, 2004, 18:05:10
Job time : 16.8883 secs

FILED DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 214:
US-09-721-456-214

Query Match 60.0%; Score 60; DB 4; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.0063;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICQIEYFGDF 18
DB 1 ICQIEYFGDF 12

RESULT 13
US-08-851-843A-24
Sequence 24, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017

FILED DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-24

Query Match 50.5%; Score 50.5; DB 3; Length 37;
Best Local Similarity 83.3%; Pred. No. 0.21;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICQIEYFGDF 18
DB 1 ICQIEYFGDF 11

RESULT 14
US-08-854-050-24
Sequence 24, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6261836el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536

APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-214

Query Match 60.0%; Score 60; DB 3; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.0063;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 7 ICEQIEYFGDF 18
||| |||||
Db 1 ICHQXEYFGDF 12

RESULT 11
US-09-402-181B-214
Sequence 214, Application US/09402181B
Patent No. 6610839
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 633
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,181B
FILING DATE: 29-Sep-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ausenhub, Scott J.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 214:
US-09-402-181B-214

Query Match 60.0%; Score 60; DB 4; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.0063;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 7 ICEQIEYFGDF 18
||| |||||
Db 1 ICHQXEYFGDF 12

RESULT 12
US-09-721-456-214
Sequence 214, Application US/09721456
Patent No. 6617110
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/721,456
FILING DATE: 22-NOV-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017

STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 215:
US-09-402-181B-215

Query Match 64.0%; Score 64; DB 4; Length 38;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICEQIEYVFGD 17
|||
1 ICEQIEYVFGD 11

Db 1 ICEQIEYVFGD 11

RESULT 9
US-09-721-456-215
Sequence 215, Application US/09721456
Patent No. 6617110

GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.

TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/721,456
FILING DATE: 22-NOV-1997
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 215:
US-09-721-456-215

Query Match 64.0%; Score 64; DB 4; Length 38;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICEQIEYVFGD 17
|||
1 ICEQIEYVFGD 11

Db 1 ICEQIEYVFGD 11

RESULT 10
US-08-974-549A-214
Sequence 214, Application US/08974549A
Patent No. 6166178

GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.

TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:

TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-25

Query Match
Best Local Similarity 100.0%; Score 64; DB 3; Length 38;
Pred. No. 0.0014;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICEQIEYFGD 17
Db 1 ICEQIEYFGD 11

RESULT 7
US-09-430-323-25
Sequence 25, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-May-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-May-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-Apr-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-Apr-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-Oct-1996

ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant

MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-430-323-25

Query Match
Best Local Similarity 100.0%; Score 64; DB 4; Length 38;
Pred. No. 0.0014;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICEQIEYFGD 17
Db 1 ICEQIEYFGD 11

RESULT 8
US-09-402-181B-215
Sequence 215, Application US/09402181B
Patent No. 6610839
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 633
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,181B
FILING DATE: 29-Sep-1997
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-Oct-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-Apr-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-Apr-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-May-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-May-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-Aug-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-Aug-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-Aug-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-Oct-1997

ATTORNEY/AGENT INFORMATION:
NAME: Ausenhus, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid

PRIOR FILING DATE: 1997-10-06
PRIOR APPLICATION NUMBER: 08/321,427
PRIOR FILING DATE: 1994-10-11
PRIOR APPLICATION NUMBER: 60/086,527
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO: 4
LENGTH: 18
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: B-LAP
OTHER INFORMATION: this peptide is biotinylated
US-09-316-630-4

Query Match 76.0%; Score 76; DB 3; Length 18;
Best Local Similarity 81.2%; Pred. No. 7.4e-06;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LDTKCEQIEYFGDF 18
Db 3 LEAKICHOIEYFGDF 18

RESULT 3
US-08-475-955-20
Sequence 20, Application US/08475955
Patent No. 6641813
GENERAL INFORMATION:
APPLICANT: Harley, John
TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS OF
TITLE OF INVENTION: AUTOANTIBODIES
NUMBER OF SEQUENCES: 218
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patricia L. Pabst
STREET: 2800 One Atlantic Center, 1250 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,955
FILING DATE: June 7, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/867,819
FILING DATE: April 13, 1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/648,205
FILING DATE: January 31, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/472,947
FILING DATE: January 31, 1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patricia L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMF114CIP(2)DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-873-8795
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Binding-site
LOCATION: 8..15
US-08-475-955-20

Query Match 65.0%; Score 65; DB 4; Length 21;
Best Local Similarity 91.7%; Pred. No. 0.00051;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ICEQIEYFGDF 18
Db 1 ICHOIEYFGDF 12

RESULT 4
US-08-851-843A-25
Sequence 25, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809e1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-25

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:43:01; Search time 15.883 Seconds
(without alignments)
58.488 Million cell updates/sec

Title: US-09-836-073-15

Perfect score: 100

Sequence: 1 LDLPKICEQIEYFGDF 18

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76.0	76.0	18	3	US-09-316-630-3
2	76.0	76.0	18	3	US-09-316-630-4
3	65.0	65.0	21	4	US-08-475-955-20
4	64.0	64.0	38	3	US-08-851-843A-25
5	64.0	64.0	38	3	US-08-974-549A-215
6	64.0	64.0	38	3	US-08-854-050-25
7	64.0	64.0	38	4	US-09-430-323-25
8	64.0	64.0	38	4	US-09-402-181B-215
9	64.0	64.0	38	4	US-09-721-456-215
10	60.0	60.0	38	3	US-08-974-549A-214
11	60.0	60.0	38	4	US-09-402-181B-214
12	60.0	60.0	38	4	US-09-721-456-214
13	50.5	50.5	37	3	US-08-851-843A-24
14	50.5	50.5	37	3	US-08-854-050-24
15	50.5	50.5	37	4	US-09-430-323-24
16	48.0	48.0	39	3	US-08-851-843A-26
17	48.0	48.0	39	3	US-08-974-549A-216
18	48.0	48.0	39	3	US-08-854-050-26
19	48.0	48.0	39	4	US-09-430-323-26
20	48.0	48.0	39	4	US-09-402-181B-216
21	48.0	48.0	39	4	US-09-721-456-216
22	44.5	44.5	178	4	US-09-107-532A-5800
23	40.0	40.0	590	1	US-08-448-196A-9
24	40.0	40.0	754	2	US-08-941-262-1
25	40.0	40.0	755	2	US-08-941-262-3
26	40.0	40.0	775	2	US-08-966-388-4
27	40.0	40.0	775	3	US-09-188-403-4

23	40	40.0	775	3	US-09-188-404-4	Sequence 4, Appl1
29	40	40.0	775	3	US-09-281-259-4	Sequence 4, Appl1
30	40	40.0	1956	3	US-08-843-417-2	Sequence 2, Appl1
31	40	40.0	1956	3	US-08-843-417-10	Sequence 10, Appl1
32	40	40.0	1956	4	US-09-527-013-2	Sequence 2, Appl1
33	40	40.0	1956	4	US-09-527-013-10	Sequence 10, Appl1
34	39	39.0	83	3	US-08-851-843A-9	Sequence 9, Appl1
35	39	39.0	83	3	US-08-851-843A-191	Sequence 191, App
36	39	39.0	83	3	US-08-854-050-9	Sequence 9, Appl1
37	39	39.0	83	4	US-09-430-323-9	Sequence 9, Appl1
38	39	39.0	83	4	US-09-402-181B-191	Sequence 191, App
39	39	39.0	83	4	US-09-721-456-191	Sequence 191, App
40	39	39.0	85	3	US-08-851-843A-11	Sequence 11, Appl
41	39	39.0	85	3	US-08-974-549A-193	Sequence 193, App
42	39	39.0	85	3	US-08-854-050-11	Sequence 11, Appl
43	39	39.0	85	4	US-09-430-323-11	Sequence 11, Appl
44	39	39.0	85	4	US-09-402-181B-193	Sequence 193, App
45	39	39.0	85	4	US-09-721-456-193	Sequence 193, App

ALIGNMENTS

RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3
Query Match 76.0%; Score 76; DB 3; Length 18;
Best local Similarity 81.2%; Pred. No. 7.4e-06;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
DB 3 LDLPKICEQIEYFGDF 18
QY 3 LDLPKICEQIEYFGDF 18
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953

This Page Blank (uspto)

OY 1 LDIDTRICEQREYFPGD 17
|||:||||| | |
DB 503 LDLEKKICEYIRSYKXD 519

Search completed: September 10, 2004, 18:00:09
Job time : 43.229 secs

```

RN SEQUENCE FROM N.A.
RP Cheuk R., Chan H., Kim C.J., Shim P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
RA Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
RA Saitoh A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinzaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis ORF clones.";
RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY050403; AAK91419.1; -
DR EMBL; BT000588; AAN18157.1; -
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:nucleus binding; IEA.
DR InterPro; IPR002344; Lupus La.
DR InterPro; IPR006630; Lupus La. dom.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR Pfam; PF00076; xtm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RM; 1.
DR PROSITE; PS0102; RM; 1.
SO SEQUENCE 422 AA; 46842 MW; 4EC4BBF1E068F0E CRC64;

Query Match 53.0%; Score 53; DB 10; Length 422;
Best Local Similarity 56.2%; Pred. No. 4.3;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 LLDTKICEQIEYFGD 17
Db 102 ELNOKIRQVEYFSD 117

RESULT 14
057230 PRELIMINARY; PRT; 1156 AA.
AC 057230;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE RNA polymerase subunit rpol32 (EC 2.7.7.6) (DNA-directed RNA
DE polymerase beta chain).
GN MVA135R.
OS Vaccinia virus (strain Ankara).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=126794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ankara;
RA Antoine G., Schefflinger F., Falkner F.G., Dörner F.;
RT "The complete genomic sequence of the Modified Vaccinia Ankara (MVA)
RT strain.";
RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N DIPHOSPHATE +
CC {RNA} (N).
DR EMBL; U94848; AAB96526.1; -
DR PIR; T37411; T37411.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003899; F:DNA-directed RNA polymerase activity; IEA.
DR GO; GO:0003900; F:DNA-directed RNA polymerase I activity; IEA.
DR GO; GO:0003901; F:DNA-directed RNA polymerase II activity; IEA.
DR GO; GO:0003902; F:DNA-directed RNA polymerase III activity; IEA.
DR GO; GO:0016740; F:transcriptase activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR007121; RNA_pol_B.
DR InterPro; IPR007645; RNA_pol_Rpb2_3.
DR InterPro; IPR007646; RNA_pol_Rpb2_4.
DR InterPro; IPR007647; RNA_pol_Rpb2_5.
DR InterPro; IPR007120; RNA_pol_Rpb2_6.
DR InterPro; IPR007641; RNA_pol_Rpb2_7.
DR Pfam; PF04565; RNA_pol_Rpb2_3; 1.
DR Pfam; PF04566; RNA_pol_Rpb2_4; 1.
DR Pfam; PF04567; RNA_pol_Rpb2_5; 1.
DR Pfam; PF04568; RNA_pol_Rpb2_6; 1.
DR Pfam; PF04569; RNA_pol_Rpb2_7; 1.
DR PROSITE; PS01166; RNA_POL_BETA; 1.
KW DNA-directed RNA polymerase; Transcription; Transferase.
SO SEQUENCE 1164 AA; 133401 MW; D28A83F6EBB8101B CRC64;

Query Match 53.0%; Score 53; DB 12; Length 1164;
Best Local Similarity 58.8%; Pred. No. 11;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

```

DR InterPro; IPR007641; RNA_pol_Rpb2_7.
DR Pfam; PF04565; RNA_pol_Rpb2_3; 1.
DR Pfam; PF04566; RNA_pol_Rpb2_4; 1.
DR Pfam; PF04567; RNA_pol_Rpb2_5; 1.
DR Pfam; PF04568; RNA_pol_Rpb2_6; 1.
DR Pfam; PF04569; RNA_pol_Rpb2_7; 1.
DR PROSITE; PS01166; RNA_POL_BETA; 1.
KW DNA-directed RNA polymerase; Transcription; Transferase.
SO SEQUENCE 1156 AA; 132425 MW; F6657C8AF5E22BC3 CRC64;

Query Match 53.0%; Score 53; DB 12; Length 1156;
Best Local Similarity 58.8%; Pred. No. 11;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LLDTKICEQIEYFGD 17
Db 495 LLDLTKICEYIRSYKD 511

RESULT 15
090027 PRELIMINARY; PRT; 1164 AA.
AC 090027;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE A25R (EC 2.7.7.6) (DNA-directed RNA polymerase beta chain).
GN A25R.
OS Variola major virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=12870;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bangladesh-1975;
RX MEDLINE=94088747; PubMed=8264798;
RA Masung R.F., Esposito J.J., Liu L., Qi J., Utechtack T.R.,
RA Knight J.C., Aubin L., Yuran T.E., Parsons J.M., Loparev V.N.,
RA Selivanov N.A., Cavaliaro K.F., Kerlavage A.R., Mahy B.W.J.,
RA Venter C.J.;
RT "Potential virulence determinants in terminal regions of variola
RT smallpox virus genome.";
RL Nature 366:748-751(1993).
CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N DIPHOSPHATE +
CC {RNA} (N).
DR EMBL; U2579; AAA60876.1; -
DR PIR; T28566; T28566.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003899; F:DNA-directed RNA polymerase activity; IEA.
DR GO; GO:0003900; F:DNA-directed RNA polymerase I activity; IEA.
DR GO; GO:0003901; F:DNA-directed RNA polymerase II activity; IEA.
DR GO; GO:0003902; F:DNA-directed RNA polymerase III activity; IEA.
DR GO; GO:0016740; F:transcriptase activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR007121; RNA_pol_B.
DR InterPro; IPR007645; RNA_pol_Rpb2_3.
DR InterPro; IPR007646; RNA_pol_Rpb2_4.
DR InterPro; IPR007647; RNA_pol_Rpb2_5.
DR InterPro; IPR007120; RNA_pol_Rpb2_6.
DR InterPro; IPR007641; RNA_pol_Rpb2_7.
DR Pfam; PF04565; RNA_pol_Rpb2_3; 1.
DR Pfam; PF04566; RNA_pol_Rpb2_4; 1.
DR Pfam; PF04567; RNA_pol_Rpb2_5; 1.
DR Pfam; PF04568; RNA_pol_Rpb2_6; 1.
DR Pfam; PF04569; RNA_pol_Rpb2_7; 1.
DR PROSITE; PS01166; RNA_POL_BETA; 1.
KW DNA-directed RNA polymerase; Transcription; Transferase.
SO SEQUENCE 1164 AA; 133401 MW; D28A83F6EBB8101B CRC64;

Query Match 53.0%; Score 53; DB 12; Length 1164;
Best Local Similarity 58.8%; Pred. No. 11;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

```

RT Staphylococcus aureus type 5 capsular polysaccharide."
RL Mol. Microbiol. 27:9-21(1998).
RN [4]
RP SEQUENCE FROM N.A.
RA STRAIN=Reynolds and Newman;
RA Baga N., Mann E.R., Foster T.J., Lee J.C.;
RU Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U81973; AAC66099.1; -.
DR HSP; P27828; 1F6D.
DR GO; GO:0008761; F:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR GO; GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
DR InterPro: IPR003331; Epimerase_2.
DR Pfam: PF02350; Epimerase_2; 1.
DR TIGRFAMs: TIGR00236; wecB; 1.
SQ SEQUENCE 391 AA; 44328 MW; B0105F690DB7CF1D CRC64;

Query Match 54.0%; Score 54; DB 2; Length 391;
Best Local Similarity 81.8%; Pred. No. 2.8;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 KICEQIEYFG 16
Db 366 RICEAIEYFG 376

RESULT 11
Q99X57 PRELIMINARY; PRT; 391 AA.
ID 099X57;
AC 099X57;
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 24, Last annotation update)
DE Capsular polysaccharide synthesis enzyme Capsb.
DE CAP OR SAV0164 OR SA0159.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OC NCB1_TaxID=158878; 158879;
OX [1]
RN SEQUENCE FROM N.A.
RP SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Koroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kubara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003358; BAB56326.1; -.
DR EMBL: AP003329; BAB41379.1; -.
DR FUR; H89777; H89777.
DR HSP; P27828; 1F6D.
DR GO; GO:0008761; F:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR GO; GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
DR InterPro: IPR003331; Epimerase_2.
DR Pfam: PF02350; Epimerase_2; 1.
DR TIGRFAMs: TIGR00236; wecB; 1.
DR Complete proteome.
SQ SEQUENCE 391 AA; 44372 MW; D0DF5FA715BCCECC CRC64;

Query Match 54.0%; Score 54; DB 16; Length 391;
Best Local Similarity 81.8%; Pred. No. 2.8;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 KICEQIEYFG 16
Db 366 RICEAIEYFG 376

```

```

RESULT 12
Q9FL36 PRELIMINARY; PRT; 411 AA.
ID Q9FL36;
AC Q9FL36;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similarity to RNA-binding protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCB1_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Columbia;
RX MEDLINE=98344145; PubMed=9679202;
RA Kaneo T., Korani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
RT features of the regions of 1,381,565 bp covered by twenty one
RT DNA Res. 5:131-145(1998)."
RL EMBL: AB010698; BAB11080.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro: IPR002344; Lupus_La.
DR InterPro: IPR006630; Lupus_La_dom.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF05383; La; 1.
DR Pfam: PF00076; rrm; 1.
DR PRINTS: PRO0302; LUPUSLA.
DR SMART: SM00715; La; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PSS0102; RRM; 1.
SQ SEQUENCE 411 AA; 45655 MW; A2EF62EB588B099 CRC64;

Query Match 53.0%; Score 53; DB 10; Length 411;
Best Local Similarity 56.2%; Pred. No. 4.2;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 DLDPTKICEQIEYFG 17
Db 91 ELNQLIRQVEYFSD 106

RESULT 13
Q94A38 PRELIMINARY; PRT; 422 AA.
ID Q94A38;
AC Q94A38;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE AT5G46250/MPL12.3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCB1_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Columbia;
RA Shum P., Chen H., Cheuk R., Kim C.-J., Koesema E., Meyers M.C.,
RA Banb J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamita A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onda C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

```

DR EMBL: AY052365; AA096556.1; -
 DR EMBL: AY139801; AA088107.1; -
 DR EMBL: AC140977; AA073903.1; -
 DR GO: GO:0003743; F:translational initiation factor activity; IEA.
 DR GO: GO:0006413; P:translational initiation; IEA.
 DR InterPro: IPR006630; Lupus la dom.
 DR InterPro: IPR002965; P:rich_extensin.
 DR InterPro: IPR001950; TIF_SUI1.
 DR Pfam: PF05383; Ia; 1.
 DR PRINTS: PR01217; PRICEXTENSIN.
 DR PROSITE: PS01118; SUI1.1; 1.
 SQ SEQUENCE 826 AA; 91377 MW; 8D41922E5B609D9A CRC64;

Query Match 55.0%; Score 55; DB 10; Length 826;
 Best Local Similarity 52.9%; Pred. No. 4;
 Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDLDTRICEQIEYFGD 17
 DB 276 LDLDTRVLRKQVEYFSD 292

RESULT 8
 O8NYS8 PRELIMINARY; PRT; 381 AA.
 AC O8NYS8;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Capsular polysaccharide synthesis enzyme Cap8P.
 GN CAP8P OR MW0139.
 OS Staphylococcus aureus (strain MW2).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=196620;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22040717; PubMed=12044378;
 RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
 RA Nagai Y., Iwano N., Asano K., Naimi T., Kuroda H., Cui L.,
 RA Yamamoto K., Hiramatsu K.;
 RT "Genome and virulence determinants of high virulence community-
 RT acquired MRSA";
 RL Lancet 359:1819-1827(2002).
 DR EMBL: AP004822; BAB94004.1; -
 DR GO: GO:0008761; F:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
 DR GO: GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
 DR GO: GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
 DR InterPro: IPR003331; Epimerase_2.
 DR Pfam: PF02350; Epimerase_2; 1.
 DR TIGRFAMs: TIGR00236; wecB; 1.
 KW Complete proteome.
 SQ SEQUENCE 381 AA; 43106 MW; 125E4D5D1904779C CRC64;

Query Match 54.0%; Score 54; DB 16; Length 381;
 Best Local Similarity 81.8%; Pred. No. 2.7;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 KICEQIEYFG 16
 DB 356 RICEAIEYFG 366

RESULT 9
 P72382 PRELIMINARY; PRT; 391 AA.
 AC P72382;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Cap8P.
 GN CAP8P.
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 RT

OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Becker;
 RX MEDLINE=96178981; PubMed=8606192;
 RA Sau S., Lee C.Y.;
 RT "Cloning of type 8 capsule genes and analysis of gene clusters for the
 RT production of different capsular polysaccharides in Staphylococcus
 RT aureus";
 RL J. Bacteriol. 178:2118-2126(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Becker;
 RX MEDLINE=97197525;
 RA Sau S., Sun J., Lee C.Y.;
 RT "Molecular characterization and transcriptional analysis of type 8
 RT capsule genes in Staphylococcus aureus";
 RL J. Bacteriol. 179:1614-1621(1997).
 DR EMBL: U73374; AAB49445.1; -
 DR HSRP: P27828; IFCD.
 DR GO: GO:0008761; F:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
 DR GO: GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
 DR GO: GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
 DR InterPro: IPR003331; Epimerase_2.
 DR Pfam: PF02350; Epimerase_2; 1.
 DR TIGRFAMs: TIGR00236; wecB; 1.
 SQ SEQUENCE 391 AA; 44247 MW; 1E8D9FA9BCT6F0D CRC64;

Query Match 54.0%; Score 54; DB 2; Length 391;
 Best Local Similarity 81.8%; Pred. No. 2.8;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 KICEQIEYFG 16
 DB 366 RICEAIEYFG 376

RESULT 10
 P95709 PRELIMINARY; PRT; 391 AA.
 AC P95709;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Cap8P.
 GN CAP8P.
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Reynolds and Newman;
 RX MEDLINE=97388587; PubMed=9245821;
 RA Sau S., Basian N., Wann E.R., Lee J.C., Foster T.J., Lee C.Y.;
 RT "The Staphylococcus aureus allelic genetic loci for serotype 5 and 8
 RT capsule expression contain the type-specific genes flanked by common
 RT genes";
 RL Microbiology 143:0-0(0).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Reynolds and Newman;
 RX MEDLINE=98101481; PubMed=9440531;
 RA Kiser K.B., Lee J.C.;
 RT "Staphylococcus aureus cap50 and cap8P genes functionally complement
 RT mutations affecting enterobacterial common-antigen biosynthesis in
 RT Escherichia coli";
 RL J. Bacteriol. 180:403-406(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Reynolds and Newman;
 RX MEDLINE=98125727; PubMed=9466251;
 RA Basian N., Albus A., Michon F., Livolsi P.J., Park J.-S., Lee J.C.;
 RT "Identification of a gene essential for O-acetylation of the

```

RESULT 5
QZ7T10 ID QZ7T10 PRELIMINARY; PRT; 401 AA.
AC 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
DE Similar to sjgren syndrome antigen B (Autocantigen Ia).
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR002344; Lupa La.
DR InterPro; IPR006630; Lupa La. dom.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 401 AA; 46138 MW; 2EF032PDD316291 CRC64;

Query Match
Best Local Similarity 64.0%; Score 64; DB 13; Length 401;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDKTICQEQIEYFGD 17
DB 10 LKKVQAEQIEYFGD 24

RESULT 6
ID 001806 PRELIMINARY; PRT; 396 AA.
AC 001806;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
DE C44B4.4 protein.
OS Caenorhabditis elegans.
OC Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Sammons L., Wohlmann P., Gillam B.;
RT "The sequence of C. elegans cosmid C44B4.";
RN Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;

```

```

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003140; AAB54169.1; -.
DR PIR; T30953; T30953.
DR WormRep; C44B4.4; CE08718.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR002344; Lupa La.
DR InterPro; IPR006630; Lupa La. dom.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF05383; Lef; 1.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
SQ SEQUENCE 396 AA; 43631 MW; A2D828A4AAAC34 CRC64;

Query Match
Best Local Similarity 58.0%; Score 58; DB 5; Length 396;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 LDKTICQEQIEYFGD 17
DB 11 DADQIKIQEQIEYFGN 26

RESULT 7
ID Q940X9 PRELIMINARY; PRT; 826 AA.
AC Q940X9;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
DE A15G21160/T10F18_190 (Prolin-rich protein family).
GN A15G21160.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shin P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Banh J., Bowser L., Carinci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinzaki K., Davis R.W., Theologis A., Becker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Shin P., Chen H., Cheuk R., Kim C.J., Bowser L., Carinci P.,
RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
RA Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinzaki K., Davis R.W., Theologis A.,
RA Becker J.R.;
RT "Arabidopsis ORF clones.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Town C.D., Haas B.J., Maiti R., Hannick L.I., Chan A.P., Roming C.M.,
RA Smith Jr R.K., Arbogast T., Tallon L.J., Utebbeck T.R., Vanaken S.E.,
RA Feldblum T.V., Yu C., Wortman J.R., White O., Fraser C.M.;
RT "Arabidopsis thaliana chromosome 5 BAC F13M11 genomic sequence.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

```

Query Match	Best Local Similarity	Score	DB	Length
Matches 11, Conservative	73.3%	68.0%	DB 13	206
Matches 11, Conservative	4	Pred. No. 0.0092	0	Indels 0, Gaps 0
Matches 11, Conservative	4	Mismatches 0	0	Indels 0, Gaps 0

This Page Blank (uspto)

```

AC 08F7G5;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Thiamine biosynthesis protein thic.
GN THIC OR LA0980.
OS Leptosira interogens.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxId=173;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J.,
RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
RA Yao Z.-L., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
RA Xu J.-G., Zhao G.-P.;
RT "Unique physiological and pathogenic features of Leptospira
RT interogens revealed by whole-genome sequencing.";
RL Nature 422:888-893 (2003).
CC -1- FUNCTION: Required for the synthesis of the hydromethylpyrimidine
CC (HMP) moiety of thiamine (4-amino-2-methyl-5-
CC hydroxymethylpyrimidine) (By similarity).
CC -1- PATHWAY: Thiamine biosynthesis.
CC -1- SIMILARITY: Belongs to the thic family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE011282; AAN48179.1; -.
CC HAMAP: MF_00089; -; 1.
CC InterPro: IPR002817; Thic.
CC Pfam: PF01964; Thic; 1.
CC ProDom: PD007048; Thic; 1.
CC TIGRfam: TIGR00190; thic; 1.
CC Thiamine biosynthesis; Complete proteome.
KW SEQUENCE 495 AA; 55989 MW; 1EB5456C99A9A87 CRC64;
SQ
Query Match 46.0%; Score 46; DB 1; Length 495;
Best Local Similarity 50.0%; Pred. No. 6.2;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 LDDPTKICEQIEY 14
DB 333 VDLQTKICQEAIFY 346

```

```

RX MEDLINE=95320184; PubMed=7597051;
RA Logemann E., Parniske M., Hahlbrock K.;
RT "Modes of expression and common structural features of the complete
RT phenylalanine ammonia-lyase gene family in parsley.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:5905-5909 (1995).
CC -1- FUNCTION: Controls carbon flux to pigments essential for
CC pollination or UV protection, to numerous phytoalexins synthesized
CC by plants when challenged by pathogens, and to lignins.
CC -1- CATALYTIC ACTIVITY: Trans-cinnamate + NADPH + O(2) = 4-
CC hydroxycinnamate + NADP(+) + H(2)O.
CC -1- PATHWAY: Lignin biosynthesis.
CC -1- PATHWAY: Phenylpropanoid metabolism; second step.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L38898; AAC11660.1; -.
CC PIR: T14907; T14907.
CC InterPro: IPR001128; Cytochrome_P450.
CC Pfam: PF00067; P450; 1.
CC PRINTS: PR00385; P450.
CC PROSITE: PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Heme; NADP.
KW METAL 448 448 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 506 AA; 58047 MW; 32F0EB959D69CCF CRC64;
QY 5 TKICEQIEYFGDF 18
DB 215 SRLAQSFYHFGDF 228

```

Search completed: September 10, 2004, 17:53:07
Job time : 8.24022 secs

DR PROSITE: PS00030: RRM RNP 1; FALSE NEG.
 KW RNA-binding; Nuclear protein; DNA-binding.
 FT DOMAIN 141 228 RNA-BINDING (RRM).
 SQ SEQUENCE 383 AA; 44430 MW; 4E5C8F21C40F452 CRC64;
 Query Match: 48.0%; Score 48; DB 1; Length 383;
 Best Local Similarity 53.3%; Pred. No. 2.2;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 LDTKICBQIEVYFGD 17
 | : |||||
 Db 43 LEASTIRQLVYFGD 57

RESULT 13
 LA DROME STANDARD; PRT; 390 AA.
 AC P40796; Q24375; Q9VIN2;
 DT 01-SEP-1995 (Rel. 31, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE La protein homolog (la ribonucleoprotein) (la autoantigen homolog).
 GN LA OR CG10922.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Canton-S; TISSUE=Ovary;
 RX MEDLINE=94309632; PubMed=8035794;
 RA Bai C., Li Z., Tolias P.P.;
 RT "Developmental characterization of a Drosophila RNA-binding protein
 RT homologous to the human systemic lupus erythematosus-associated
 RT La/SS-B autoantigen.";
 RL Mol. Cell. Biol. 14:5123-5129(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94309661; PubMed=8035818;
 RA Yoo C.J., Molin S.L.;
 RT "La proteins from Drosophila melanogaster and Saccharomyces
 RT cerevisiae: a yeast homolog of the la autoantigen is dispensable for
 RT growth.";
 RL Mol. Cell. Biol. 14:5412-5424(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.B., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.W., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fodor C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalaal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laake P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milne N.V., Modarres C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,
 RA Palazzolo M., Plutman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: May be involved in transcription termination by RNA
 CC polymerase III. Binds RNA and DNA. Binds to precursors of RNA
 CC polymerase III transcripts. May play a specialized role during fly
 CC development.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DEVELOPMENTAL STAGE: Expressed throughout embryonic, larval,
 CC pupal, and adult development. Expression throughout the embryo is
 CC followed by a restricted pattern of mesodermal expression that is
 CC later confined to the visceral mesoderm, gonads, gut, and salivary
 CC glands.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC DR EMBL: U07652; AAA20518.1; -;
 CC DR EMBL: L32988; AAA21776.1; -;
 CC DR EMBL: AE003666; AA853885.1; -;
 CC DR PIR: A53773; A53773.
 CC DR PIR: A53781; A53781.
 CC DR FLYBASE: FBgn0011638; La.
 CC DR GO: GO:0008098; F:5S rRNA primary transcript binding; IDA.
 CC DR GO: GO:0003723; F:RNA binding; NAS.
 CC DR InterPro: IPR002344; Lupus_La.
 CC DR InterPro: IPR006630; Lupus_La_dom.
 CC DR InterPro: IPR00504; RNA_rec_mot.
 CC DR Pfam: PF05383; La; 1.
 CC DR Pfam: PF00076; rtm; 1.
 CC DR PRINTS: PR00302; LUPUSLA.
 CC DR SMART: SM00715; LA; 1.
 CC DR SMART: SM00360; RRM; 1.
 CC DR PROSITE: PS5102; RRM; 1.
 CC DR PROSITE: PS00030; RRM RNP 1; 1.
 CC KW RNA-binding; Nuclear protein; DNA-binding.
 CC FT DOMAIN 149 234
 CC FT CONFLICT 169 169 A -> T (IN REF. 1).
 CC FT CONFLICT 182 183 KH -> NS (IN REF. 1).
 CC FT CONFLICT 283 283 A -> R (IN REF. 1).
 CC FT CONFLICT 329 329 K -> N (IN REF. 1).
 CC SQ SEQUENCE 390 AA; 44884 MW; A80992889044655 CRC64;

Query Match: 48.0%; Score 48; DB 1; Length 390;
 Best Local Similarity 72.7%; Pred. No. 2.2;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICBOIEVYFGD 17
 | : |||||
 Db 54 IIRQVYVFGD 64

RESULT 14
 THIC_LBPIN STANDARD; PRT; 495 AA.

```

CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M57687; AAA67907.1; -.
CC DR EMBL; L22858; AAA66725.1; -.
CC DR PIR; A38499; HJNVAV.
CC DR InterPro; IPR006824; Baculo_helicase.
CC DR Pfam; PF04735; Baculo_helicase.1.
CC DR Helicase; DNA replication; ATP-binding; DNA-binding; Nuclear protein;
CC Early protein.
CC FT DOMAIN 692 701 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC FT NP_BIND 917 924 ATP (BY SIMILARITY).
CC FT DNA_BIND 967 981 H-T-H MOTIF (BY SIMILARITY).
CC FT MUTAGEN 934 934 V->M: DEFECTIVE IN THE SYNTHESIS OF VIRAL
CC SHUTOFF OF HOST PROTEIN SYNTHESIS AND THE
CC NONPERMISSIVE TEMPERATURE (MUTANT T58).
CC FT CONFLICT 126 126 F -> S (IN REF. 1).
CC FT CONFLICT 1149 1149 F -> L (IN REF. 1).
CC SQ SEQUENCE 1221 AA; 143213 MW; 090E199855862D1B CRC64;

Query Match 51.0%; Score 51; DB 1; Length 1221;
Best Local Similarity 61.5%; Pred. No. 2.4;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 DTKICEQIEYRFG 16
Db 440 DTKICVSLGYRFG 452

RESULT 11
IPR MYCPU STANDARD; PRT; 183 AA.
AC Q98096;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-
DE hydrolase) (Ppase).
GN PPA OR MYPV 4700.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxId=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAB CTIP;
RX MEDLINE=21267165; PubMed=11353084;
RA Chandaud I., Heilig R., Ferris S., Barbe V., Samson D., Gallison F.,
RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,
RA Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
RT Mycoplasma pulmonis."
RT Nucleic Acids Res. 29:2145-2153(2001).
CC -1- CATALYTIC ACTIVITY: diphosphate + H(2)O = 2 phosphate.
CC -1- COFACTOR: Binds 4 magnesium ions per subunit. Other metal ions can
CC support activity, but at a lower rate. Two magnesium ions are
CC required for the activation of the enzyme and are present before
CC substrate binds, two additional magnesium ions form complexes with
CC substrate and product (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the Ppase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL445564; CAC13643.1; -.
CC DR PIR; F90570; F90570.
CC DR Mypulstat; MYPV_4700; -.
CC DR HAMAP; MF_00209; -.
CC DR InterPro; IPR008163; Inorg_ppnaph.
CC DR InterPro; IPR008162; Pyrophosphatase.
CC DR Pfam; PF00719; Pyrophosphatase.1.
CC DR PRODOM; PD002014; Inorg_ppnaph.1.
CC DR PROSITE; PS00387; PPA5E.1.
CC DR Hydrolase; Metal-binding; Magnesium; Complete proteome.
CC FT METAL 53 53 MAGNESIUM 1 (BY SIMILARITY).
CC FT METAL 58 58 MAGNESIUM 1 AND 2 (BY SIMILARITY).
CC FT METAL 90 90 MAGNESIUM 1 (BY SIMILARITY).
CC SQ SEQUENCE 183 AA; 21715 MW; 65AF81A063AAAD3 CRC64;

Query Match 50.0%; Score 50; DB 1; Length 183;
Best Local Similarity 47.1%; Pred. No. 0.47;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 2 DTDKICEQIEYRFGDF 18
Db 111 DDDSKWLEIKYFSGNY 127

RESULT 12
LA AEDAL STANDARD; PRT; 383 AA.
AC Q26457;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE A protein homolog (La ribonucleoprotein) (La autoantigen homolog).
DE Aedes albopictus (Forest day mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.
OX NCBI_TaxId=7160;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96135233; PubMed=8551578;
RA Parignon N., Strauss J.H.;
RT "Mosquito homolog of the La autoantigen binds to Sindbis virus RNA."
RT J. Virol. 70:1173-1181(1996).
CC -1- FUNCTION: May be involved in transcription termination by RNA
CC polymerase III. Binds RNA and DNA. Binds to the 3' end of the
CC minus strand of Sindbis virus RNA. This may be significant for
CC Sindbis virus RNA replication.
CC -1- SUBCELLULAR LOCATION: Nuclear. Primarily nuclear, but significant
CC amounts are present in the cytoplasm.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S80954; AAB35931.1; -.
CC DR InterPro; IPR002344; Lupaus_La.
CC DR InterPro; IPR006630; Lupaus_La_dom.
CC DR InterPro; IPR000504; RNA_rec_mot.
CC DR Pfam; PF05383; La.1.
CC DR Pfam; PF00076; Irm.1.
CC DR PRINTS; PR00302; LUPUSLA.
CC DR SMART; SM00715; LA.1.
CC DR SMART; SM00360; RRM.1.
CC DR PROSITE; PS50102; RRM.1.

```

RT "Appendix to 'The complete DNA sequence of vaccinia virus'";
RL Virology 179:517-563(1990).
CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
CC of DNA into RNA using the four ribonucleoside triphosphates as
CC substrates.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -1- SUBUNIT: This enzyme consists of at least eight subunits.
CC -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M37415; AAA72882.1; -;
DR EMBL; M35027; AAA48148.1; -;
DR PIR; H42519; RNZ87.
DR InterPro; IPR007121; RNA_pol_B.
DR InterPro; IPR007645; RNA_pol_Rpb2_3.
DR InterPro; IPR007646; RNA_pol_Rpb2_4.
DR InterPro; IPR007647; RNA_pol_Rpb2_5.
DR InterPro; IPR007120; RNA_pol_Rpb2_6.
DR InterPro; IPR007641; RNA_pol_Rpb2_7.
DR Pfam; PF04565; RNA_pol_Rpb2_3; 1.
DR Pfam; PF04566; RNA_pol_Rpb2_4; 1.
DR Pfam; PF04567; RNA_pol_Rpb2_5; 1.
DR Pfam; PF00562; RNA_pol_Rpb2_6; 1.
DR Pfam; PF04560; RNA_pol_Rpb2_7; 1.
DR PROSITE; PS01166; RNA_POL_BETA; 1.
DR Transferrase; DNA-directed RNA polymerase; Transcription; Zinc;
KW Zinc-finger.
FT ZN FING 1087 1106 C4-TYPE (POTENTIAL).
SQ SEQUENCE 1164 AA; 133363 MW; B1F5C1484BA37D0D CRC64;
Qy Query Match 53.0%; Score 53; DB 1; Length 1164;
Best Local Similarity 58.8%; Pred. No. 1.1;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Db 1 LDLDTKICEQIEYYFGD 17
503 LDLEKXICEYIRSYKD 519
RESULT 9
RPO2_VARV STANDARD; PRT; 1164 AA.
AC P33811;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE DNA-directed RNA polymerase 132 kDa polypeptide (EC 2.7.7.6).
GN RPO132 OR A24R.
OS Variola virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10255;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=India-1967 / Isolate Ind3;
RX MEDLINE=93202281; PubMed=8384129;
RA Shchelkunov S.N., Blinov V.M., Sandakhchiev L.S.;
RT "Genes of variola and vaccinia viruses necessary to overcome the host
RT protective mechanisms";
RL FEBS Lett. 319:80-83(1993).
CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
CC of DNA into RNA using the four ribonucleoside triphosphates as
CC substrates.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).

CC -1- SUBUNIT: This enzyme consists of at least eight subunits.
CC -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X69198; CAA49069.1; -;
DR PIR; G36850; G36850.
DR InterPro; IPR007121; RNA_pol_B.
DR InterPro; IPR007645; RNA_pol_Rpb2_3.
DR InterPro; IPR007646; RNA_pol_Rpb2_4.
DR InterPro; IPR007647; RNA_pol_Rpb2_5.
DR InterPro; IPR007120; RNA_pol_Rpb2_6.
DR InterPro; IPR007641; RNA_pol_Rpb2_7.
DR Pfam; PF04565; RNA_pol_Rpb2_3; 1.
DR Pfam; PF04566; RNA_pol_Rpb2_4; 1.
DR Pfam; PF04567; RNA_pol_Rpb2_5; 1.
DR Pfam; PF00562; RNA_pol_Rpb2_6; 1.
DR Pfam; PF04560; RNA_pol_Rpb2_7; 1.
DR PROSITE; PS01166; RNA_POL_BETA; 1.
DR Transferrase; DNA-directed RNA polymerase; Transcription; Zinc;
KW Zinc-finger.
FT ZN FING 1087 1106 C4-TYPE (POTENTIAL).
SQ SEQUENCE 1164 AA; 133329 MW; 0452B84ED810CD53 CRC64;
Qy Query Match 53.0%; Score 53; DB 1; Length 1164;
Best Local Similarity 58.8%; Pred. No. 1.1;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Db 1 LDLDTKICEQIEYYFGD 17
503 LDLEKXICEYIRSYKD 519
RESULT 10
V143_NPVAC STANDARD; PRT; 1221 AA.
AC P24307;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Helicase.
GN P143.
OS Autographa californica nuclear polyhedrosis virus (AcMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=46015;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HR3;
RX MEDLINE=91134998; PubMed=1994581;
RA Lu A., Carstens E.B.;
RT "Nucleotide sequence of a gene essential for viral DNA replication in
RT the baculovirus Autographa californica nuclear polyhedrosis virus";
RL Virology 181:336-347(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C6;
RX MEDLINE=94303173; PubMed=8030224;
RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
RT "The complete DNA sequence of Autographa californica nuclear
RT polyhedrosis virus";
RL Virology 202:586-605(1994).
CC -1- FUNCTION: Essential for the initiation of viral DNA replication,
CC it may contribute to other functions such as controlling the
CC switch to the late phase and leading to the inhibition of host
CC protein synthesis. Required for late and very late gene
CC expression.

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: X67859; CAA48043.1; -
 CC PIR: JCI1494; JCI1494.
 CC InterPro: IPR002344; Lupus_La.
 CC InterPro: IPR006630; Lupus_La_dom.
 CC InterPro: IPR00504; RNA_rec_mot.
 CC Pfam: PF05383; La; 1.
 CC Pfam: PF00076; rrm; 1.
 CC PRINTS: PR00302; LUPUSLA.
 CC SMART: SM00715; LA; 1.
 CC SMART: SM00360; RRM; 1.
 CC PROSITE: PS50102; RRM; 1.
 CC PROSITE: PS00030; RRM_RNP_1; 1.
 CC RNA-binding; Nuclear protein; Phosphorylation.
 CC DOMAIN 111 187
 CC SEQUENCE 415 AA; 47777 MW; 033FD9C1E475F98 CRC64;

Query Match 76.0%; Score 76; DB 1; Length 415;
 Best Local Similarity 81.2%; Pred. No. 5.7e-05;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 LDTKICEQIEYFGDF 18
 DB 13 LEAKICHOIEYFGDF 28

RESULT 7
 RPO2_COMPX STANDARD; PRT; 1164 AA.
 AC P17474; Q90025;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE DNA-directed RNA polymerase 132-kDa polypeptide (EC 2.7.7.6).
 GN RPO132.
 OS CompoX virus (CPV).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OX NCBI_TaxID=10243;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Brighton red;
 RX MEDLINE=89125698; PubMed=2915377;
 RA Patel D.D., Pickup D.J.;
 RT "The second-largest subunit of the poxvirus RNA polymerase is similar
 RT to the corresponding subunits of procaryotic and eucaryotic RNA
 RT polymerases.";
 RL J. Virol. 63:1076-1086(1989).
 CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
 CC of DNA into RNA using the four ribonucleoside triphosphates as
 CC substrates.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA} (n).
 CC -1- SUBUNIT: This enzyme consists of at least eight subunits.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative initiation;
 CC Comment=2 isoforms, 1/late specific (shown here) and 2/early and
 CC late, are produced by alternative initiation. One
 CC transcriotional start site is operative at late times only and
 CC the other is operative both at early and late times;
 CC -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: M26173; AAA42919.1; -
 CC EMBL: M26173; AAA42920.1; -
 CC PIR: A31879; RNVCZP.
 CC InterPro: IPR007121; RNA_pol_B.
 CC InterPro: IPR007645; RNA_pol_Rpb2_3.
 CC InterPro: IPR007646; RNA_pol_Rpb2_4.
 CC InterPro: IPR007647; RNA_pol_Rpb2_5.
 CC InterPro: IPR007120; RNA_pol_Rpb2_6.
 CC InterPro: IPR007641; RNA_pol_Rpb2_7.
 CC Pfam: PF04565; RNA_pol_Rpb2_3; 1.
 CC Pfam: PF04566; RNA_pol_Rpb2_4; 1.
 CC Pfam: PF04567; RNA_pol_Rpb2_5; 1.
 CC Pfam: PF00562; RNA_pol_Rpb2_6; 1.
 CC Pfam: PF04560; RNA_pol_Rpb2_7; 1.
 CC PROSITE: PS01166; RNA_POL_BETA; 1.
 CC -transferase; DNA-directed RNA polymerase; Transcription; Zinc;
 CC Zinc-finger; Alternative initiation.
 CC CHAIN 1 1164
 CC CHAIN 1 1164
 CC FT CHAIN 9 1164
 CC FT INIT MET 9 1106
 CC FT ZN FING 1087 1106
 CC FT SEQUENCE 1164 AA; 133323 MW; 220D6EF5238DBF7 CRC64;

Query Match 53.0%; Score 53; DB 1; Length 1164;
 Best Local Similarity 58.8%; Pred. No. 1.1;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 LDTKICEQIEYFGDF 17
 DB 503 LDEKKICEYRSYKDF 519

RESULT 8
 RPO2_VACCV STANDARD; PRT; 1164 AA.
 AC P19758;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA-directed RNA polymerase 132 kDa polypeptide (EC 2.7.7.6).
 GN RPO132 OR A24R.
 OS Vaccinia virus (strain WR), and
 OS Vaccinia virus (strain Copenhagen).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OX NCBI_TaxID=10254; 10249;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WR;
 RX MEDLINE=91082452; PubMed=1824607;
 RA Amesgazi B.Y., Holmes M.H., Cole N.B., Jones E.V., Earl P.L.,
 RA Moss B.;
 RT "Identification, sequence, and expression of the gene encoding the
 RT second-largest subunit of the vaccinia virus DNA-dependent RNA
 RT polymerase.";
 RL Virology 180:88-98(1991).
 CC [2]
 CC SEQUENCE FROM N.A.
 CC STRAIN=Copenhagen;
 CC MEDLINE=91021027; PubMed=2219722;
 CC Goebel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,
 CC Paoletti E.;
 CC "The complete DNA sequence of vaccinia virus.";
 CC Virology 179:247-266(1990).
 CC [3]
 CC COMPLETE GENOME.
 CC STRAIN=Copenhagen;
 CC Goebel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,
 CC Paoletti E.;

FT DOMAIN 111 187 RNA-BINDING (RRM).
 FT MOD RES 366 366 PHOSPHORYLATION (BY CK2).
 SQ SEQUENCE 408 AA; 46837 MW; EC153C15F9187FC4 CRC64;
 Query Match 76.0%; Score 76; DB 1; Length 408;
 Best Local Similarity 81.2%; Pred. No. 5.6e-05;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYFGDF 18
 Db 13 LEAKICHQIEYFGDF 28

RESULT 5
 LA_MOUSE STANDARD; PRT; 415 AA.
 ID LA_MOUSE
 AC P32067;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen homolog)
 DE SSB OR SS-B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93203630; PubMed=8454877;
 RA Topfer F., Gordon T., McCluskey J.;
 RT "Characterization of the mouse autoantigen La (SS-B). Identification of conserved RNA-binding motifs, a putative ATP binding site and reactivity of recombinant protein with poly(U) and human autoantibodies.";
 RT J. Immunol. 150:3091-3100(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary gland;
 RX MEDLINE=23288257; PubMed=1247932;
 RA Klausner R.L., Feigold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
 RA Stopleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rabe S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Boeck S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.U., Lu X., Gibbs R.A.,
 RA Fahy J., Helton B., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield Y.S.N., Krzywinski M.T., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE OF 1-11 FROM N.A.
 RA Grodz D., Bachmann M.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' terminus of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S, and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: Phosphorylated (By similarity).
 CC -1- PTM: Phosphorylated (By similarity).

CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; L00993; AAA39415.1; -
 CC EMBL; BC003820; AA03820.1; -
 CC EMBL; Y07951; CAA69249.1; -
 CC MGD; MGI:98423; Ssb.
 CC GO; GO:0005634; C:nucleus; IDA.
 CC InterPro; IPR002344; Lupus_La.
 CC InterPro; IPR006630; Lupus_La_dom.
 CC InterPro; IPR000504; RNA_rec_mot.
 CC Pfam; PF05383; La; 1.
 CC Pfam; PF00076; rrm; 1.
 CC PRINTS; PR00302; LUPUSLA.
 CC SMART; SM00715; LA; 1.
 CC SMART; SM00360; RRM; 1.
 CC PROSITE; PS50102; RRM; 1.
 CC PROSITE; PS00030; RRM_RNP_1; 1.
 KW RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 111 187
 SQ SEQUENCE 415 AA; 47756 MW; 2D75197692FDC933 CRC64;
 Query Match 76.0%; Score 76; DB 1; Length 415;
 Best Local Similarity 81.2%; Pred. No. 5.7e-05;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYFGDF 18
 Db 13 LEAKICHQIEYFGDF 28

RESULT 6
 LA_MOUSE STANDARD; PRT; 415 AA.
 ID LA_MOUSE
 AC P36656;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen homolog)
 DE SSB OR SS-B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93246255; PubMed=7916708;
 RA Senses I., Troester H., Bartsch H., Schemmle M., Igloi G.L.,
 RA Bachmann M.;
 RT "Isolation of rat cDNA clones coding for the autoantigen SS-B/La: detection of species-specific variations.";
 RT Gene 126:265-268(1993).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' terminus of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S, and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: Phosphorylated (By similarity).
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

Query Match 76.0%; Score 76; DB 1; Length 404;
 Best Local Similarity 81.2%; Pred. No. 5.5e-05;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDITCEQIEYFGDF 18
 | : ||| ||| ||| |||
 Db 13 LEAKICHOIEYFGDF 28

RESULT 4
 LA_HUMAN STANDARD; PRT; 408 AA.
 ID P05455;
 AC 01-NOV-1988 (Rel. 09, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein (Sjogren syndrome type B antigen) (SS-B) (La
 DE ribonucleoprotein) (La autoantigen).
 GN SSB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 OX 1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89202037; PubMed=2468131;
 RA Chan E.K.L., Sullivan K.F., Tan E.M.;
 RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus
 RT Nucleic Acids Res. 17:2233-2244(1989).
 RL [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=89053970; PubMed=3192525;
 RA Chambers J.C., Kenan D., Martin B.J., Keene J.D.;
 RT "Genomic structure and amino acid sequence domains of the human La
 RT autoantigen."
 RL J. Biol. Chem. 263:18043-18051(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta, and Skeletal muscle;
 RX MEDLINE=22388257; PubMed=1247932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Cohen F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Ziegler B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein U.E., Jones S.U.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE OF 54-408 FROM N.A.
 RX MEDLINE=88199081; PubMed=2452201;
 RA Sturge S.A.D., Peterson M.G., McNeillage L.J., Whittingham S.,
 RA Coppel R.S.;
 RT "Characteristics and epitope mapping of a cloned human autoantigen
 RT La."
 RL J. Immunol. 140:3212-3218(1988).
 RN [5]
 RP SEQUENCE OF 54-97 FROM N.A.
 RX MEDLINE=8516283; PubMed=8556886;
 RA Chambers J.C., Keene J.D.;

RT "Isolation and analysis of cDNA clones expressing human lupus La
 RT antigen."
 RT Proc. Natl. Acad. Sci. U.S.A. 82:2115-2119(1985).
 RL [6]
 RN FUNCTION.
 RX MEDLINE=89251617; PubMed=2470590;
 RA Gottlieb E., Stelz J.A.;
 RT "Function of the mammalian La protein: evidence for its action in
 RT transcription termination by RNA polymerase III."
 RL EMBO J. 8:851-861(1989).
 RN [7]
 RP PHOSPHORYLATION.
 RX MEDLINE=97207017; PubMed=9054510;
 RA Fan H., Sakulich A.L., Goodier J.L., Zhang X., Qin J., Marais R.J.;
 RT "Phosphorylation of the human La antigen on serine 366 can regulate
 RT recycling of RNA polymerase III transcription complexes."
 RL Cell 88:707-715(1997).
 RN [8]
 RP INTERACTION WITH DDX15.
 RX MEDLINE=22346609; PubMed=12458796;
 RA Fournaux M.A., Kolman M.J.M., Van der Heijden A., De Jong A.S.,
 RA Van Venrooij W.J., Pruijn G.J.M.;
 RT "The human La (SS-B) autoantigen interacts with DDX15/hPrp43, a
 RT putative DEAD-box RNA helicase."
 RL RNA 8:1428-1443(2002).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' terminus of virtually all nascent
 CC polymerase III transcripts. It is associated with precursor forms
 CC of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S,
 CC and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE
 CC C-TERMINAL PART OF THE PROTEIN.
 CC -1- PTM: The N-terminus is blocked.
 CC -1- DISASS: Sera from patients with systemic lupus erythematosus
 CC often contain antibodies that react with the normal cellular
 CC La protein as if this antigen was foreign.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X13697; CAA31985.1; -
 CC EMBL; J04205; AAA51885.1; -
 CC EMBL; BC001289; AAH01289.1; -
 CC EMBL; BC020818; AAH20818.1; -
 CC PIR; A31888; A31888.
 CC Genew; HGNC:11316; SSB.
 CC MIM; 109090; -
 CC GO; GO:0030529; C:ribonucleoprotein complex; TAS.
 CC GO; GO:0003729; F:mRNA binding; TAS.
 CC GO; GO:0000049; F:RNA binding; TAS.
 CC GO; GO:0008334; P:histone mRNA metabolism; TAS.
 CC GO; GO:0006400; P:RNA modification; TAS.
 CC InterPro; IPR002344; Lupus_La.
 CC InterPro; IPR006630; Lupus_La_dom.
 CC InterPro; IPR005054; RNA_rec_mot.
 CC Pfam; PF00383; La; 1.
 CC Pfam; PF00076; rrm; 1.
 CC PRINTS; PRO0302; LUPUSLA.
 CC SMART; SM00715; LA; 1.
 CC SMART; SM00360; RRM; 1.
 CC PROSITE; PS50102; RRM_1.
 CC PROSITE; PS00030; RRM_RNP_1; 1.
 CC Systemic lupus erythematosus; RNA-binding; Phosphorylation;
 CC Nuclear protein.

DR PROSITE: PS00030; RRM_RNP_1; FALSE NEG.
 KM RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 110 202 RNA-BINDING (RRM).
 FT DOMAIN 315 331 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 SQ SEQUENCE 427 AA; 48995 MW; 45F3146F8934A355 CRC64;

Query March 94.0%; Score 94; DB 1; Length 427;
 Best Local Similarity 100.0%; Pred. No. 6.2e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LDDPTKICEQIEYFGD 17
 |||||
 Db 10 LDDPTKICEQIEYFGD 26

RESULT 2
 ID LAA_XENLA STANDARD; PRT; 428 AA.
 AC P28048;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lopus La protein homolog A (La ribonucleoprotein A) (La autoantigen homolog A).
 GN LAA1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocyte;
 RX MEDLINE=93287095; PubMed=8510143;
 RA Scherly D., Stutz F., Lin-Marg N., Clarkson S.G.;
 RT "La proteins from Xenopus laevis. cDNA cloning and developmental expression.";
 RL J. Mol. Biol. 231:196-204(1993).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' terminus of virtually all nascent polymerase III transcripts (by similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DEVELOPMENTAL STAGE: Barely detectable in stage I/II oocytes, accumulate in stage III/IV oocytes, then exhibit a roughly constant steady state level in mature oocytes, eggs, and early embryos.
 CC -1- PTM: Phosphorylated (Probable).
 CC -1- MISCELLANEOUS: There are two forms of La, Laa and Lab, in xenopus.
 CC -1- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, X68817; CAA48715.1; -.
 CC PIR, S33818; S33818.
 CC InterPro; IPR002344; Lopus_La.
 CC InterPro; IPR006630; Lopus_La_dom.
 CC InterPro; IPR000504; RNA_rec_mot.
 CC Pfam; PF05383; La; 1.
 CC Pfam; PF00076; rtm; 1.
 CC PRINTS; PR00302; LUPUSLA.
 CC SMART; SM00715; LA; 1.
 CC SMART; SM00360; RRM; 1.
 CC PROSITE; PS50102; RRM; 1.
 CC PROSITE; PS00030; RRM_RNP_1; 1.
 CC RNA-binding; Nuclear protein; Phosphorylation.

FT DOMAIN 111 203 RNA-BINDING (RRM).
 FT DOMAIN 196 212 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 316 332 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 SQ SEQUENCE 428 AA; 48864 MW; AEB3A38B7D2E3EC3 CRC64;

Query March 88.0%; Score 88; DB 1; Length 428;
 Best Local Similarity 94.1%; Pred. No. 6.1e-07;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LDDPTKICEQIEYFGD 17
 |||||
 Db 11 LDDPTKICEQIEYFGD 27

RESULT 3
 ID LA_BOVIN STANDARD; PRT; 404 AA.
 AC P10881;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lopus La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 GN SSB.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pituitary;
 RX MEDLINE=89202037; PubMed=2468131;
 RA Chan E.K.L., Sullivan K.F., Tan E.M.;
 RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences for RNA-binding.";
 RL Nucleic Acids Res. 17:2233-2244(1989).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' terminus of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including crRNA and 4.5S, 5S, 7S, and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15 (by similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE C-TERMINAL PART OF THE PROTEIN.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, X13698; CAA31986.1; -.
 CC PIR, S03849; S03849.
 CC InterPro; IPR002344; Lopus_La.
 CC InterPro; IPR006630; Lopus_La_dom.
 CC InterPro; IPR000504; RNA_rec_mot.
 CC Pfam; PF05383; La; 1.
 CC Pfam; PF00076; rtm; 1.
 CC PRINTS; PR00302; LUPUSLA.
 CC SMART; SM00715; LA; 1.
 CC SMART; SM00360; RRM; 1.
 CC PROSITE; PS50102; RRM; 1.
 CC PROSITE; PS00030; RRM_RNP_1; 1.
 CC RNA-binding; Nuclear protein; Phosphorylation.
 CC DOMAIN 111 187 RNA-BINDING (RRM).
 SQ SEQUENCE 404 AA; 46534 MW; 4EB30B5C262AD6A1 CRC64;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:01:41; Search time 7.24022 Seconds
(without alignments)
129.452 Million cell updates/sec

Title: US-09-836-073-15

Perfect score: 100

Sequence: 1 LDLDTKICEQIFRYRGDF 18

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94	94.0	427	1	LAB_XENLA
2	88	88.0	428	1	LAB_XENLA
3	76	76.0	404	1	LA_BOVIN
4	76	76.0	408	1	LA_HUMAN
5	76	76.0	415	1	LA_MOUSE
6	76	76.0	415	1	LA_RAT
7	53	53.0	1164	1	RPO2_COMPLEX
8	53	53.0	1164	1	RPO2_VACCV
9	53	53.0	1164	1	RPO2_VARV
10	51	51.0	1221	1	V143_NPVAC
11	50	50.0	183	1	IPYR_MYCPU
12	48	48.0	390	1	LA_AEDAL
13	48	48.0	390	1	LA_DROME
14	46	46.0	495	1	THIC_LERIN
15	45	45.0	506	1	TCMO_PECRC
16	44	44.0	490	1	IFET4_HUMAN
17	44	44.0	1069	1	PCHT_HUMAN
18	43	43.0	366	1	DHPH_THRIN
19	43	43.0	711	1	ETP2_MXYVL
20	43	43.0	711	1	ETP2_SFYKA
21	43	43.0	2184	1	POLG_ECOLI
22	42	42.0	270	1	YC87_METUA
23	41	41.0	298	1	LAH1_SCHPO
24	41	41.0	438	1	2277_HUMAN
25	41	41.0	1104	1	IPYR_HAERIN
26	40	40.0	176	1	IPYR_HAERIN
27	40	40.0	247	1	ADCI_RHILLO
28	40	40.0	256	1	ADCS_RHILLO
29	40	40.0	261	1	ADCC_RHILLO
30	40	40.0	466	1	SRO9_YEAST
31	40	40.0	474	1	SYE_VIBPA
32	40	40.0	552	1	FET4_YEAST
33	40	40.0	608	1	ALB1_SALSA

34	40	40.0	608	1	ALB2_SALSA	003156 salmo salar
35	40	40.0	755	1	SEC6_RAT	062825 rattus norv
36	40	40.0	756	1	SEC6_HUMAN	060645 homo sapien
37	40	40.0	2911	1	FBN2_HUMAN	P35556 homo sapien
38	39.5	39.5	297	1	YX01_CAEL	011108 caenorhabdi
39	39.5	39.5	859	1	ST7_HUMAN	095651 homo sapien
40	39.5	39.5	967	1	SYL_PYRHO	058698 pyrococcus
41	39	39.0	175	1	IPYR_PSEPK	088916 pseudomonas
42	39	39.0	191	1	Y948_METUA	058358 methanococc
43	39	39.0	245	1	YCEL_BACST	045401 bacillus st
44	39	39.0	264	1	GRAX_HUMAN	P49863 homo sapien
45	39	39.0	305	1	HEM6_VIBCH	09kvt4 vibrio chol

ALIGNMENTS

RESULT 1
LAB_XENLA STANDARD; PRT; 427 AA.
AC P28049;
DT 01-AUG-1992 (rel. 23, Created)
DT 01-AUG-1992 (rel. 23, Last sequence update)
DT 15-MAR-2004 (rel. 43, Last annotation update)
DE lupus La protein homolog B (La ribonucleoprotein B) (La autoantigen homolog B).
GN LAB1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
XP [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RX MEDLINE=93287095; PubMed=8510143;
RA Scherly D., Stultz F., Lin-Marg N., Clarkson S.G.;
RT "La proteins from Xenopus laevis. cDNA cloning and developmental expression.";
RT J. Mol. Biol. 231:196-204 (1993).
RL
CC -|- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' terminus of virtually all nascent polymerase III transcripts (By similarity).
CC -|- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -|- DEVELOPMENTAL STAGE: Barely detectable in stage I/II oocytes, accumulate in stage III/IV oocytes, then exhibit a roughly constant steady state level in mature oocytes, eggs, and early embryos.
CC -|- PTM: Phosphorylated (Probable).
CC -|- MISCELLANEOUS: There are two forms of La, Laa and Lab, in xenopus.
CC -|- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
CC -|- SIMILARITY: Contains 1 RNA recognition motif (RNM) domain.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL: X68818; CAA48716.1; -
CC PIR: S33817; S33817.
CC InterPro: IPR002344; Lupus_La.
CC InterPro: IPR006630; Lupus_La_dom.
CC InterPro: IPR00504; RNA_rec_mot.
CC 3fam; PF05383; La; 1.
CC 3fam; PF00076; rrm; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00715; LA; 1.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PSS0102; RRM; 1.

This Page Blank (uspto)

A:Reference number: A72850; MUID:94303173; PMID:8030224

A:Accession: H72861

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-125, 'F', 127-1148, 'F', 1150-1221 <AYR>

A:Cross-references: GB:L22858; MID:G510708; PIDN:AAA6725.1; PID:G559164

C:Genetics:

A:Gene: Ac-helicase

C:Superfamily: AcMNPV helicase

C:Keywords: ATP; DNA binding; DNA repair; DNA replication; hydrolase; nucleotide binding

F:917-924/Region: nucleotide-binding motif A (P-loop)

Query Match 51.0%; Score 51; DB 1; Length 1221;

Best Local Similarity 61.5%; Pred. No. 6.9; Mismatches 3; Indels 0; Gaps 0;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 DTRICEQIEYFPG 16

Db 440 DTRLCVSLGYFG 452

RESULT 15

F90570

hypothetical protein MYPV 4700 [Imported] - Mycoplasma pulmonis (strain UAB CTIP)

C:Species: Mycoplasma pulmonis

C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001

C:Accession: F90570

R:Chabaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;

Nucleic Acids Res. 29, 2145-2153, 2001

A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm

A:Reference number: A99512; MUID:21267165; PMID:11353084

A:Accession: F90570

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-183 <KUR>

A:Cross-references: GB:AL445566; PID:G14089884; PIDN:CAC13643.1; GSPDB:GN00153

A:Experimental source: strain UAB CTIP

C:Genetics:

A:Gene: MYPV 4700

A:Genetic code: SGC3

C:Superfamily: inorganic pyrophosphatase

Query Match 50.0%; Score 50; DB 2; Length 183;

Best Local Similarity 47.1%; Pred. No. 1.4; Mismatches 3; Indels 0; Gaps 0;

Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 DLDTRICEQIEYFPG 18

Db 111 DLDKWLLEIKYFSSNY 127

Search completed: September 10, 2004, 18:02:33
Job time : 16.0782 secs

A:Title: The complete DNA sequence of vaccinia virus.
 A:Reference number: A42531; MUID:91021027; PMID:2219722
 A:Contents: annotation; possible protein-coding frames
 A>Note: neither amino acid nor nucleotide sequence is given
 C:Superfamily: DNA-directed RNA polymerase 132K polypeptide
 C:Keywords: leucine zipper; nucleotidyltransferase; transcription; zinc finger
 F:1087-1106/Region: zinc finger CCCC motif

Query Match 53.0%; Score 53; DB 1; Length 1164;
 Best Local Similarity 58.8%; Pred. No. 3.1;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LDLDTKICEQIEYFGD 17
 |||:|||||:|
 Db 503 LDLEKKICEYIRSYKD 519

RESULT 10

RNVZCP
 DNA-directed RNA polymerase (EC 2.7.7.6) 132K chain - cowpox virus (strain Brighton Red)
 C:Species: cowpox virus
 C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 11-Jun-1999
 C:Accession: A31879
 R:Patel, D.D.; Pickup, D.J.
 J:Virology 63, 1076-1086, 1989

A:Title: The second-largest subunit of the poxvirus RNA polymerase is similar to the core
 A:Reference number: A31879; MUID:89125698; PMID:2915377
 A:Accession: A31879
 A:Molecule type: DNA
 A:Residues: 1-1164 <PART>
 A:Cross-references: GB:M26173; NID:G323393; PIDN:AAA42919.1; PID:G323394
 C:Comment: This enzyme consists of at least seven subunits whose molecular weights are 1

A:Gene: rpo132
 C:Superfamily: DNA-directed RNA polymerase 132K polypeptide
 C:Keywords: nucleotidyltransferase; transcription; zinc finger

Query Match 53.0%; Score 53; DB 1; Length 1164;
 Best Local Similarity 58.8%; Pred. No. 3.1;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LDLDTKICEQIEYFGD 17
 |||:|||||:|
 Db 503 LDLEKKICEYIRSYKD 519

RESULT 11

T28566
 DNA-directed RNA polymerase (EC 2.7.7.6) 133K chain - variola major virus
 C:Species: variola major virus
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
 C:Accession: T28566
 R:Maesung, R.F.; Bapostito, J.J.; Liu, L.I.; Qi, J.; Uteerback, T.R.; Knight, J.C.; Aubin

Nature 366, 748-751, 1993
 A:Title: Potential virulence determinants in terminal regions of variola smallpox virus
 A:Reference number: Z20488; MUID:94088747; PMID:8264798
 A:Accession: T28566
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1164 <MS>

A:Cross-references: EMBL:L22579; NID:G623595; PIDN:AAA60876.1; PID:G439046
 A:Experimental source: strain Bangladesh-1975
 C:Superfamily: DNA-directed RNA polymerase 132K polypeptide
 C:Keywords: nucleotidyltransferase; transcription

Query Match 53.0%; Score 53; DB 2; Length 1164;
 Best Local Similarity 58.8%; Pred. No. 3.1;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LDLDTKICEQIEYFGD 17
 |||:|||||:|
 Db 503 LDLEKKICEYIRSYKD 519

RESULT 12

F72166
 A25R protein - variola minor virus (strain Garcia-1966)

C:Species: variola minor virus

C:Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 20-Jun-2000

C:Accession: F72166

R:Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Safonov, P.F.; Maesung, R.F.; Lopez

A:Description: Analysis of the complete coding sequence of DNA of a strain variola minor

A:Reference number: A72150

A:Accession: F72166

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1164 <SHC>

A:Cross-references: GB:Y16780; NID:G5830555; PIDN:CAB54728.1; PID:G5830689

A:Experimental source: strain Garcia-1966

C:Genetics:

A:Gene: A25R

C:Superfamily: DNA-directed RNA polymerase 132K polypeptide

Query Match 53.0%; Score 53; DB 2; Length 1164;
 Best Local Similarity 58.8%; Pred. No. 3.1;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LDLDTKICEQIEYFGD 17
 |||:|||||:|
 Db 503 LDLEKKICEYIRSYKD 519

RESULT 13

G36850
 A24R protein - variola virus (strain India-1967)

C:Species: variola virus

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 23-Mar-2001

C:Accession: G36850

R:Bilimov, V.M.

submitted to GenBank, November 1992

A:Reference number: A36859

A:Accession: G36850

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1164 <BLI>

A:Cross-references: GB:X69198; NID:G456758; PIDN:CAA49069.1; PID:G297307

C:Superfamily: DNA-directed RNA polymerase 132K polypeptide

Query Match 53.0%; Score 53; DB 2; Length 1164;
 Best Local Similarity 58.8%; Pred. No. 3.1;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LDLDTKICEQIEYFGD 17
 |||:|||||:|
 Db 503 LDLEKKICEYIRSYKD 519

RESULT 14

H0N1AV

helicase (EC 3.6.1.-) - Autographa californica nuclear polyhedrosis virus

C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 19-Jan-2001

C:Accession: A38499; H72861

R:Lu, A.; Carstens, E.B.

Virology 181, 336-347, 1991

A:Title: Nucleotide sequence of a gene essential for viral DNA replication in the baculo

A:Reference number: A38499; MUID:91134998; PMID:1994581

A:Accession: A38499

A:Molecule type: DNA

A:Residues: 1-1221 <LUA>

A:Cross-references: EMBL:M57687

R:Ayres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Farber, M.; Posee, R.D.

Virology 202, 586-605, 1994

A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.

Gene 126, 265-268, 1993
 A>Title: Isolation of rat cDNA clones coding for the autoantigen SS-B/La: Detection of A
 A/Reference number: JCI494; MUID:93246255; PMID:7916708
 A/Accession: JCI494
 A/Molecule type: mRNA
 A/Residues: 1-415 <SEM>
 A/Cross-references: GB:X67859; NID:G55778; PIDN:CAA48043.1; PID:G55779
 A/Experimental source: liver
 A/Comment: This protein associates with a variety of small RNA molecules, most of which
 act as a transcription termination factor.
 C/Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C/Keywords: phosphoprotein; RNA binding
 F.112-178/Domain: ribonucleoprotein repeat homology <RRM>
 F.113-118/Region: RNA-binding RNP2 motif
 F.151-158/Region: RNA-binding RNP1 motif
 F.127-415/Domain: phosphorylated #status predicted <P>

Query Match 76.0%; Score 76; DB 1; Length 415;
 Best Local Similarity 81.2%; Pred. No. 0.00021;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYFFGDF 18
 DB 13 LEAKICHQIEYFFGDF 28

RESULT 6
 T30953
 hypothetical protein C4E4.4 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000
 C/Accession: T30953
 R/Sammons, L.; Wohldmann, P.; Gilliam, B.
 submitted to the EMBL Data Library, August 1999
 A/Description: The sequence of C. elegans cosmid C4E4.
 A/Reference number: Z20945
 A/Accession: T30953
 A/Status: preliminary; translated from GB/EMBL/DDBJ
 A/Molecule type: DNA
 A/Residues: 1-396 <SAM>
 A/Cross-references: EMBL:AF003140; PIDN:AB54169.1
 A/Experimental source: strain Bristol N2; clone C4E4
 C/Genetics:
 A/Map position: 1
 A/Introns: 45/1; 114/3
 A/Note: C4E4.4
 C/Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

Query Match 58.0%; Score 58; DB 2; Length 396;
 Best Local Similarity 62.5%; Pred. No. 0.16;
 Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 DLDTKICEQIEYFFGDF 17
 DB 11 DADQKIKKQLEFFGN 26

RESULT 7
 H89777
 capsular polysaccharide synthesis enzyme CapSP [imported] - Staphylococcus aureus (strain
 C/Species: Staphylococcus aureus
 C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C/Accession: H89777
 R/Auruda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A/Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
 A/Reference number: A89758; MUID:21311952; PMID:11418146
 A/Accession: H89777
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-391 <KTR>

A/Cross-references: GB:BA000018; PID:G13700080; PIDN:BAE41379.1; GSPDB:GN00149
 A/Experimental source: strain N315
 C/Genetics:
 A/Gene: capP
 C/Superfamily: lipopolysaccharide biosynthesis protein bplD

Query Match 54.0%; Score 54; DB 2; Length 391;
 Best Local Similarity 81.8%; Pred. No. 0.68;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 KICEQIEYFFG 16
 DB 366 RICEAIEYFFG 376

RESULT 8
 T37411
 RNA polymerase subunit rpol32 - vaccinia virus (strain Ankara)
 C/Species: vaccinia virus
 A/Variety: strain Ankara
 C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000
 C/Accession: T37411
 R/Antoine, G.; Scheiflinger, F.; Falkner, F.G.; Dörner, F.
 submitted to the EMBL Data Library, March 1997
 A/Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain
 A/Reference number: Z20877
 A/Accession: T37411
 A/Status: preliminary; translated from GB/EMBL/DDBJ
 A/Molecule type: DNA
 A/Residues: 1-1156 <ANT>
 A/Cross-references: EMBL:U94848; PIDN:AB96526.1
 A/Experimental source: strain Ankara
 C/Genetics:
 A/Note: MVA135R
 C/Superfamily: DNA-directed RNA polymerase 132K polypeptide

Query Match 53.0%; Score 53; DB 2; Length 1156;
 Best Local Similarity 58.8%; Pred. No. 3.1;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LDDTKICEQIEYFFGDF 17
 DB 495 LDLEKKICEYRSYND 511

RESULT 9
 RNV28T
 DNA-directed RNA polymerase (EC 2.7.7.6) 132K chain - vaccinia virus
 N/Alternate names: A24R protein
 C/Species: vaccinia virus
 A/Note: host Homo sapiens (man)
 C/Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 11-Jun-1999
 C/Accession: H42519; A38517
 R/Goebel, S.J.; Johnson, G.P.; Parkus, M.E.; Davis, S.W.; Winslow, J.P.; Paolletti, E.
 Virology 179, 517-563, 1990
 A/Title: Appendix to "The complete DNA sequence of vaccinia virus".
 A/Reference number: A42501
 A/Accession: H42519
 A/Molecule type: DNA
 A/Residues: 1-1164 <GOE>
 A/Cross-references: GB:M35027; NID:G335317; PIDN:AAA48148.1; PID:G335496
 A/Experimental source: strain Copenhagen
 R/Amegadzie, B.Y.; Holmes, M.H.; Cole, N.B.; Jones, E.V.; Earl, P.L.; Moss, B.
 Virology 180, 88-98, 1991
 A/Title: Identification, sequence, and expression of the gene encoding the second-large
 A/Reference number: A38517; MUID:91082452; PMID:1824607
 A/Accession: A38517
 A/Molecule type: DNA
 A/Residues: 1-1164 <AME>
 A/Cross-references: EMBL:M37415; NID:G335800; PIDN:AAA72882.1; PID:G335801
 A/Experimental source: strain WR
 R/Goebel, S.J.; Johnson, G.P.; Parkus, M.E.; Davis, S.W.; Winslow, J.P.; Paolletti, E.
 Virology 179, 247-266, 1990

F.112-178/Domain: ribonucleoprotein repeat homology <RNM>
 F.113-118/Region: RNA-binding RNP2 motif
 F.151-158/Region: RNA-binding RNP1 motif
 F.128-428/Domain: phosphorylated #status predicted <PHY>

Query Match 88.0%; Score 88; DB 1; Length 428;
 Best Local Similarity 94.1%; Pred. No. 2.5e-06;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LDDTKICEQIEYFGDF 17
 11 LDDTKICEQIEYFGDF 27

DB 11 LDDTKICEQIEYFGDF 27

RESULT 3

S03849
 ribonucleoprotein la - bovine
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 31-Dec-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S03849
 R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.
 Nucleic Acids Res. 17, 2233-2244, 1989
 A>Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences
 A:Reference number: S03848; MUID:89202037; PMID:2468131
 A:Accession: S03849
 A:Molecule type: mRNA
 A:Residues: 1-404 <CHA>
 A:Cross-references: EMBL:X13698; NID:g755; PIDN:CAA31986.1; PID:g756
 A:Note: Part of this sequence was confirmed by protein sequencing
 C:Comment: This protein associates with a variety of small RNA molecules, most of which
 act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology
 C:Keywords: blocked amino end; phosphoprotein; RNA binding
 F.112-178/Domain: ribonucleoprotein repeat homology <RNM>
 F.113-118/Region: RNA-binding RNP2 motif
 F.151-158/Region: RNA-binding RNP1 motif
 F.128-404/Domain: phosphorylated #status predicted <PHY>

Query Match 76.0%; Score 76; DB 1; Length 404;
 Best Local Similarity 81.2%; Pred. No. 0.0002;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 LDTKICEQIEYFGDF 18
 13 LEAKICHOIEYFGDF 28

DB 13 LEAKICHOIEYFGDF 28

RESULT 4

A31888
 ribonucleoprotein la - human
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B; Sjogren syndrome antigen
 C:Species: Homo sapiens (man)
 C>Date: 21-May-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: A31888; S03848; J22956; A61051; S11013; IS5553; I70205; I70206; A31273
 R:Chambers, J.C.; Kennan, D.; Martin, B.J.; Keene, J.D.
 J. Biol. Chem. 263, 18043-18051, 1988
 A>Title: Genomic structure and amino acid sequence domains of the human la autoantigen.
 A:Reference number: A31888; MUID:89053970; PMID:3192525
 A:Accession: A31888
 A:Molecule type: mRNA
 A:Residues: 1-408 <CHA>
 A:Cross-references: GB:J04205; NID:g178686; PIDN:AAA51885.1; PID:g178687
 R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.
 Nucleic Acids Res. 17, 2233-2244, 1989
 A>Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences
 A:Reference number: S03848; MUID:89202037; PMID:2468131
 A:Accession: S03848
 A:Molecule type: mRNA
 A:Residues: 1-408 <CHA>
 A:Cross-references: EMBL:X13697; NID:g36414; PIDN:CAA31985.1; PID:g36415
 R:Chambers, J.C.; Keene, J.D.
 Proc. Natl. Acad. Sci. U.S.A. 82, 2115-2119, 1985

A>Title: Isolation and analysis of cDNA clones expressing human lupus la antigen.
 A:Reference number: A22956; MUID:85166283; PMID:3856888
 A:Accession: A22956
 A:Molecule type: mRNA
 A:Residues: 45-97, 'LK' <CH3>
 A:Cross-references: GB:J04205
 A:Note: this sequence has been revised in reference A31888
 R:Nyman, U.; Ringertz, N.R.; Pettersson, I.
 Immunol. Lett. 22, 65-72, 1989
 A>Title: Demonstration of an amino terminal la epitope recognized by human anti-la sera.
 A:Reference number: A61051; MUID:89379261; PMID:2476379
 A:Accession: A61051
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-19, 'E', 21-47 <NYM>
 R:Sturges, A.D.; Peterson, M.G.; McNeilage, L.J.; Whittingham, S.; Coppel, R.L.
 J. Immunol. 140, 3212-3218, 1988
 A>Title: Characteristics and epitope mapping of a cloned human autoantigen la.
 A:Reference number: S11013; MUID:88199081; PMID:2452201
 A:Accession: S11013
 A:Molecule type: mRNA
 A:Residues: 'E', 55-287, 'V', 289-408 <STU>
 A:Cross-references: EMBL:M20328; NID:g337456; PIDN:AAA36577.1; PID:g337457
 R:Kohsaka, H.; Yamamoto, K.; Fujii, H.; Mura, H.; Miyasaka, N.; Nishioaka, K.; Miyamoto, J.
 Clin. Invest. 85, 1566-1574, 1990
 A>Title: Fine epitope mapping the human SS-B/La protein: Identification of a distinct au
 A:Reference number: I55553; MUID:90237237; PMID:1692037
 A:Accession: I55553
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 81-107 <RES>
 A:Cross-references: GB:M5261; NID:g338491; PIDN:AAA36652.1; PID:g338495
 A:Accession: I70205
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 174-224 <RE2>
 A:Cross-references: GB:M5263; NID:g338492; PIDN:AAA36653.1; PID:g338496
 A:Accession: I70206
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 279-342 <RE3>
 A:Cross-references: GB:M5262; NID:g338493; PIDN:AAA36654.1; PID:g338497
 C:Comment: This protein associates with a variety of small RNA molecules, most of which
 act as a transcription termination factor.
 C:Genetics:
 A:Gene: GDB:SSB
 A:Cross-references: GDB:125359; OMIM:109090
 A:Map position: 2
 A:Introns: 22/3; 57/2; 115/3; 185/2; 209/2; 223/3; 264/3; 380/2
 C:Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F.112-178/Domain: ribonucleoprotein repeat homology <RNM>
 F.113-118/Region: RNA-binding RNP2 motif
 F.151-158/Region: RNA-binding RNP1 motif
 F.128-408/Domain: phosphorylated #status experimental <PHY>

Query Match 76.0%; Score 76; DB 1; Length 408;
 Best Local Similarity 81.2%; Pred. No. 0.0002;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 LDTKICEQIEYFGDF 18
 13 LEAKICHOIEYFGDF 28

DB 13 LEAKICHOIEYFGDF 28

RESULT 5

J01494
 ribonucleoprotein la - rat
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 24-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: J01494; S25145
 R:Semsel, I.; Troester, H.; Bartsch, H.; Schwemmler, M.; Igloi, G.L.; Bachmann, M.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:33:11; Search time 14.0782 Seconds
(without alignments)
122.988 Million cell updates/sec

Title: US-09-836-073-15

Perfect score: 100

Sequence: 1.LDLDTKICEQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94.0	427	1	S33817	ribonucleoprotein
2	88.0	428	1	S33818	ribonucleoprotein
3	76.0	404	1	S03849	ribonucleoprotein
4	76.0	403	1	A31888	ribonucleoprotein
5	76.0	415	1	JC1494	ribonucleoprotein
6	58.0	396	2	T30953	hypothetical prote
7	54.0	391	2	H89777	capsular polysacch
8	53.0	1156	2	T37411	RNA polymerase sub
9	53.0	1164	1	RNVZBT	DNA-directed RNA p
10	53.0	1164	1	RNVZCP	DNA-directed RNA p
11	53.0	1164	2	T28566	DNA-directed RNA p
12	53.0	1164	2	F72166	A25R protein - var
13	53.0	1164	2	G36850	A24R protein - var
14	51.0	1221	1	HUNNAV	helicase (EC 3.6.1
15	50.0	183	2	F90570	hypothetical prote
16	48.0	190	2	A47569	p143 DNA helicase
17	48.0	390	2	A53773	la/SS-B homolog D-
18	48.0	390	2	A53781	ribonucleoprotein
19	48.0	1222	2	T41835	DNA helicase P143
20	47.0	166	2	T32701	hypothetical prote
21	47.0	529	2	T00677	hypothetical prote
22	45.0	399	2	T01035	hypothetical prote
23	45.0	308	2	T14907	trans-cinnamate 4-
24	44.0	165	2	T29244	hypothetical prote
25	44.0	1069	2	T00043	BH-protocadherin-a
26	44.0	1072	2	T00041	BH-protocadherin p
27	44.0	1200	2	T00042	BH-protocadherin p
28	44.0	1743	2	T26859	hypothetical prote
29	43.0	292	2	T02349	hypothetical prote

30	43.0	366	2	J00513	phenylalanine dehy
31	43.0	505	2	B90181	Na+/H+ antiporter
32	43.0	658	2	D96566	hypothetical prote
33	43.0	759	2	AC0368	probable autotrans
34	43.0	788	2	A71076	hypothetical prote
35	42.0	199	2	A11512	weakly methyltrans
36	42.0	270	2	F64460	hypothetical prote
37	42.0	277	2	AD0121	probable exported
38	42.0	421	2	AE2473	hypothetical prote
39	42.0	1055	2	AD2499	hypothetical prote
40	41.5	140	2	C96957	hypothetical prote
41	41.0	63	2	C96942	hypothetical prote
42	41.0	199	2	B97180	probable Zn-depend
43	41.0	260	2	H71979	probable type II r
44	41.0	298	2	T18937	rna binding protei
45	41.0	298	2	T43542	RNA-binding protei

ALIGNMENTS

RESULT 1
S33817
ribonucleoprotein la.B - African Clawed frog
N:Alternate names: autoantigen SS-B/la.; ribonucleoprotein SS-B
C:Species: Xenopus laevis (African clawed frog)
C>Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C:Accession: S33817; S28544
J. Mol. Biol. 231, 196-204, 1993
A>Title: la proteins from Xenopus laevis. cDNA cloning and developmental expression.
A:Reference number: S33817; MUID:93287095; PMID:8510143
A:Accession: S33817
A:Molecule type: mRNA
A:Residues: 1-427 <SCH>
A:Cross-references: EMBL:X68818; NID:G64875; PIDD:CAA48716.1; PIDD:G64876
C:Comment: This protein associates with a variety of small RNA molecules, most of which
act as a transcription termination factor.
C:Keywords: ribonucleoprotein la; ribonucleoprotein repeat homology
F:111-177/Domain: ribonucleoprotein repeat homology <RHM>
F:112-117/Region: RNA-binding RNP2 motif
F:150-157/Region: RNA-binding RNP1 motif
F:227-427/Domain: phosphorylated #status predicted <PHY>

Query Match 94.0%; Score 94; DB 1; Length 427;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LDLDTKICEQIEYFGD 17
DB 10 LDLDTKICEQIEYFGD 26

RESULT 2
S33818
ribonucleoprotein la.A - African Clawed frog
N:Alternate names: autoantigen SS-B/la.; ribonucleoprotein SS-B
C:Species: Xenopus laevis (African clawed frog)
C>Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
R:Sherry, D.; Stutz, F.; Lin-Marg, N.; Clarkson, S.G.
J. Mol. Biol. 231, 196-204, 1993
A>Title: la proteins from Xenopus laevis. cDNA cloning and developmental expression.
A:Reference number: S33817; MUID:93287095; PMID:8510143
A:Accession: S33818
A:Molecule type: mRNA
A:Residues: 1-428 <SCH>
A:Cross-references: EMBL:X68817; NID:G64873; PIDD:CAA48715.1; PIDD:G64874
C:Comment: This protein associates with a variety of small RNA molecules, most of which
act as a transcription termination factor.
C:Keywords: ribonucleoprotein la; ribonucleoprotein repeat homology

DT 01-OCT-2000 (TReMBLrel. 15, last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
DE Similar to KIA00731 protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eubacteriobacteriota; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT clone.p0499C11.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP001080; BAA90356.1; -.
DR Gramene; Q9LJ02; -.
DR InterPro; IPR006607; DUF DM15.
DR InterPro; IPR006630; Lupus_La_dom.
DR Pfam; PF05383; La; 1.
DR SMART; SM00684; DM15; 3.
DR SMART; SM00715; La; 1.
SQ SEQUENCE 928 AA; 103745 MW; 53FC46E24A446B4 CRC64;

Query Match 49.0%; Score 48.5; DB 10; Length 928;
Best Local Similarity 68.8%; Pred. No. 25;
Matches 11; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

OY 3 LEAKICHIQIEYF-GD 17
DB 278 LRAKILTOVEYFSGD 293

RESULT 15
O44678 PRELIMINARY; PRT; 91 AA.
ID O44678
AC O44678;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein C14C6.12.
GN C14C6.12.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA MEDLINE=99069613; PubMed=9851916;
RX Wilson R.;
RA "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RT Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA David M., Wohlmann P., Bauer C., Antoniou B.;
RT "The sequence of C. elegans cosmid C14C6.";
RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF039051; AAB94258.2; -.
KW Hypothetical protein.
SQ SEQUENCE 91 AA; 10407 MW; D16D2BD4011F8762 CRC64;

Query Match 48.5%; Score 48; DB 5; Length 91;
Best Local Similarity 69.2%; Pred. No. 3;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 6 KICHIQIEYFSGD 18
DB 6 KICNVIETMISDF 18

Search completed: September 10, 2004, 18:00:07
Job time : 42.229 secs

```

DE CapsP.
CN CAPSP.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Reynolds and Newman;
RX MEDLINE=9738587; PubMed=9245821;
RA "Sau S., Bhasin N., Mann E.R., Lee J.C., Foster T.J., Lee C.Y.;
RT "The Staphylococcus aureus allelic genetic loci for serotype 5 and 8
RT capsule expression contain the type-specific genes flanked by common
RT genes."
RL Microbiology 143:0-0(0).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN-Reynolds and Newman;
RX MEDLINE=98101481; PubMed=9440531;
RA Kiser K.B., Lee J.C.;
RT "Staphylococcus aureus caps5 and capsP genes functionally complement
RT mutations affecting enterobacterial common-antigen biosynthesis in
RT Escherichia coli."
RL J. Bacteriol. 180:403-406(1998).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN-Reynolds and Newman;
RX MEDLINE=98125727; PubMed=9466251;
RA Bhasin N., Albun A., Michon F., Livolsi P.J., Park J.-S., Lee J.C.;
RT "Identification of a gene essential for O-acetylation of the
RT Staphylococcus aureus type 5 capsular polysaccharide."
RL Mol. Microbiol. 27:9-21(1998).
RN [4]
RN SEQUENCE FROM N.A.
RC STRAIN-Reynolds and Newman;
RX Bagg N., Mann E.R., Foster T.J., Lee J.C.;
RT Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL: U81973; AAC46099.1; -.
DR HSSP; P27828; 1F6D.
DR GO; GO:0008761; F:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR GO; GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
DR InterPro; IPR003331; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.
DR TIGRPFAM; TIGR00236; wecB; 1.
SQ SEQUENCE 391 AA; 44328 MW; B0105F690DB7CF1D CRC64;

Query Match 49.5%; Score 49; DB 2; Length 391;
Best Local Similarity 72.7%; Pred. No. 8.7;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIEYFG 16
DB 366 RICEAIEYFG 376

RESULT 12
Q99X57 PRELIMINARY; PRT; 391 AA.
AC O99X57;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Capsular polysaccharide synthesis enzyme CapsP.
GN CAP OR SAV0164 OR SA0159.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OC Staphylococcus aureus (strain Mu50).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RN SEQUENCE FROM N.A.
RC SEQUES=.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311957; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,

```

```

RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kaneshita M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hatford M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003358; BAB56326.1; -.
DR EMBL; AP003129; BAB41379.1; -.
DR PIR; H89777; H89777.
DR HSSP; P27828; 1F6D.
DR GO; GO:0008761; F:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR GO; GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
DR InterPro; IPR003331; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.
DR TIGRPFAM; TIGR00236; wecB; 1.
KW Complete proteome.
SQ SEQUENCE 391 AA; 44372 MW; D0DF5FA715BCCECC CRC64;

Query Match 49.5%; Score 49; DB 16; Length 391;
Best Local Similarity 72.7%; Pred. No. 8.7;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIEYFG 16
DB 366 RICEAIEYFG 376

RESULT 13
Q8LMP9 PRELIMINARY; PRT; 481 AA.
ID Q8LMP9;
AC Q8LMP9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OSJNBA001114.12.
OS Oryza sativa (japonica cultivar-group).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Eukaryota; Viridiplantae; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriaristaceae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RN SEQUENCE FROM N.A.
RA Wang R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RA Sasaki C., Henry D., Oates R., Simmons J.;
RT "Rice Genomic Sequence."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC105730; AAM51838.1; -.
DR Gramene; Q8LMP9; -.
DR InterPro; IPR006630; Lupus_La_dom.
DR Pfam; PF05383; La; 1.
DR SMART; SM00715; LA; 1.
KW Hypothetical protein.
SQ SEQUENCE 481 AA; 49959 MW; 6EA2BED311191362 CRC64;

Query Match 49.5%; Score 49; DB 10; Length 481;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 LEAKICHOIEYFG 17
DB 307 LRAELKCHIEYFSD 321

RESULT 14
Q9LJ02 PRELIMINARY; PRT; 928 AA.
ID Q9LJ02;
AC Q9LJ02;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)

```

OC SpERMATOPHYTES; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 OK NCBI_TaxID=3702;
 [1]
 RP SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
 RA Bahr J., Bowser L., Carinici P., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M.,
 RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RA "Arabidopsis cDNA clones."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carinici P.,
 RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
 RA Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
 RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
 RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
 RA Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RA "Arabidopsis ORF clones."
 RT Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AY050403; AAK91419.1; -;
 DR EMBL; BT000588; AAN18157.1; -;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 SQ SEQUENCE 422 AA; 46842 MW; 4EC4BBBF1E068F0E CRC64;
 Query Match 50.5%; Score 50; DB 10; Length 422;
 Best Local Similarity 60.0%; Pred. No. 6.4;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 Oy 3 LEAKICHOIEYFGD 17
 Db 103 LNKITROVEYFSD 117
 RESULT 9
 OBNYN8 PRELIMINARY; PRT; 381 AA.
 AC OBNYN8;
 DT 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Capsular polysaccharide synthesis enzyme Cap8P.
 CN CAP8P OR MW0139.
 OS Staphylococcus aureus (strain MW2).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=196620;
 OX [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=22040717; PubMed=12044378;
 RA Bahr J., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
 RA Nagai Y., Iwano N., Asano K., Naimi T., Kuroda H., Cui L.,
 RA Yamamoto K., Hiramatsu K.;
 RA "Genome and virulence determinants of high virulence community-
 RT acquired MRSA."
 RT Lancet 359:1819-1827(2002).
 RL EMBL; AP004822; BAB94004.1; -;

DR GO; GO:0008761; F:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
 DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
 DR GO; GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
 DR InterPro; IPR003331; Epimerase_2.
 DR Pfam; PF02350; Epimerase_2; 1.
 DR TIGRFAMs; TIGR00236; wecB; 1.
 KW Complete proteome.
 SQ SEQUENCE 381 AA; 43106 MW; 125E4D5D1904779C CRC64;
 Query Match 49.5%; Score 49; DB 16; Length 381;
 Best Local Similarity 72.7%; Pred. No. 8.5;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Oy 6 KICHOIEYFG 16
 Db 366 RICEAIEYFG 366
 RESULT 10
 ID P72382 PRELIMINARY; PRT; 391 AA.
 AC P72382;
 DT 01-FEB-1997 (TRENBLrel. 02, Created)
 DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Cap8P.
 DN CAP8P.
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Becker;
 RA MEDLINE=96178981; PubMed=8606192;
 RA Sau S., Lee C.Y.;
 RT "Cloning of type 8 capsule genes and analysis of gene clusters for the
 RT production of different capsular polysaccharides in Staphylococcus
 RT aureus."
 RL J. Bacteriol. 178:2118-2126(1996).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Becker;
 RX MEDLINE=97197525;
 RA Sau S., Sun J., Lee C.Y.;
 RT "Molecular characterization and transcriptional analysis of type 8
 RT capsule genes in Staphylococcus aureus."
 RL J. Bacteriol. 179:1614-1621(1997).
 DR EMBL; U73374; AAB49445.1; -;
 DR HSBP; P27828; IFED.
 DR GO; GO:0008761; F:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
 DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
 DR GO; GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
 DR InterPro; IPR003331; Epimerase_2.
 DR Pfam; PF02350; Epimerase_2; 1.
 DR TIGRFAMs; TIGR00236; wecB; 1.
 SQ SEQUENCE 391 AA; 42427 MW; 1E8D9FA93BC76F0D CRC64;
 Query Match 49.5%; Score 49; DB 2; Length 391;
 Best Local Similarity 72.7%; Pred. No. 8.7;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Oy 6 KICHOIEYFG 16
 Db 366 RICEAIEYFG 376
 RESULT 11
 ID P95709 PRELIMINARY; PRT; 391 AA.
 AC P95709;
 DT 01-MAY-1997 (TRENBLrel. 03, Created)
 DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)

```

RESULT 5
QZRTIO PRELIMINARY; PRT; 401 AA.
ID QZRTIO
AC QZRTIO;
DT 01-JUN-2003 (TEMBLrel. 24, Created)
DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Similar to sjogren syndrome antigen B (Autoantigen La).
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC045392; AAH45392.1; -.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR InterPro: IPR002344; Lupus_La.
DR InterPro: IPR006630; Lupus_La_dom.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 1.
DR PRINTS: PR00302; LUPUSLA.
DR SMART: SM00715; LA; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS50102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; 1.
DR PROSITE: PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 401 AA; 46138 MW; 2EF032FDD3916291 CRC64;

Query Match
Best Local Similarity 73.3%; Score 62; DB 13; Length 401;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LEAKICHOIEYFGD 17
DB 10 LEKVAIEYFGD 24

RESULT 6
QRT8V5 PRELIMINARY; PRT; 390 AA.
ID QRT8V5
AC QRT8V5;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE AT22034P.
DE LA OR CG10922.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Flise E.,
RA George R., Gonzalez M., Guatin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Paclel J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celinker S.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY075257; AA68124.1; -.
DR FLYbase: FBgn0011638; La.
DR GO: GO:0008098; F:5S rRNA primary transcript binding; IDA.
DR GO: GO:0003723; F:RNA binding; NAS.
DR InterPro: IPR002344; Lupus_La.
DR InterPro: IPR006630; Lupus_La_dom.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF05383; La; 1.

```

```

DR Pfam: PF00076; rrm; 1.
DR PRINTS: PR00302; LUPUSLA.
DR SMART: SM00715; LA; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS50102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; 1.
DR PROSITE: PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 390 AA; 44869 MW; 797FDE26B903C909 CRC64;

Query Match
Best Local Similarity 51.5%; Score 51; DB 5; Length 390;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGD 17
DB 51 ERAIRQVEYFGD 64

RESULT 7
Q9FL36 PRELIMINARY; PRT; 411 AA.
ID Q9FL36
AC Q9FL36;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Similarity to RNA-binding protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Euroside II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=98344145; PubMed=9679202;
RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
RT features of the regions of 1,381,565 bp covered by twenty one
RT physically assigned pl and TAC clones.";
RL DNA Res. 5:131-145 (1998).
DR EMBL: AB010698; BAB11080.1; -.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR InterPro: IPR002344; Lupus_La.
DR InterPro: IPR006630; Lupus_La_dom.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF05383; La; 1.
DR Pfam: PF00076; rrm; 1.
DR PRINTS: PR00302; LUPUSLA.
DR SMART: SM00715; LA; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS50102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 411 AA; 45655 MW; A2EF62EB589B099 CRC64;

Query Match
Best Local Similarity 60.0%; Score 50; DB 10; Length 411;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 LEAKICHOIEYFGD 17
DB 92 LNKIRQVEYFGD 106

RESULT 8
Q9AA38 PRELIMINARY; PRT; 422 AA.
ID Q9AA38
AC Q9AA38;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE AT5G46250/MP12.3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

DR GO:0005634; C:nucleus; IDA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR005054; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
SQ SEQUENCE 381 AA; 43891 MW; 2E2DEF1452C0F0E9 CRC64;

Query Match
Best Local Similarity 96.0%; Score 95; DB 11; Length 381;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALEAKICHOIEYFGDF 18
Db 12 ALEAKICHOIEYFGDF 28

RESULT 2
Q8BTU4 PRELIMINARY; PRT; 415 AA.
ID Q8BTU4;
AC Q8BTU4;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Sjogren syndrome antigen B.
GN SSB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOB; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK088677; BAC40498.1; -.
DR MGD; MGI:98423; Seb.
DR GO; GO:0005634; C:nucleus; IDA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR005054; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM_1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 415 AA; 47657 MW; A7545C7686AC6363 CRC64;

Query Match
Best Local Similarity 96.0%; Score 95; DB 11; Length 415;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALEAKICHOIEYFGDF 18
Db 12 ALEAKICHOIEYFGDF 28

RESULT 3
Q7ZTK2 PRELIMINARY; PRT; 427 AA.
ID Q7ZTK2;
AC Q7ZTK2;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Similar to lupus LA protein homolog B.

```

```

OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC046554; AA046554.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR005054; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 427 AA; 48996 MW; 1E7CD82D8AB9C69A CRC64;

Query Match
Best Local Similarity 70.7%; Score 70; DB 13; Length 427;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LEAKICHOIEYFGD 17
Db 12 LDTKICHOIEYFGD 26

RESULT 4
Q8OH15 PRELIMINARY; PRT; 206 AA.
ID Q8OH15;
AC Q8OH15;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Untranslated region binding-protein.
GN UBP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA L'Ecuylver T.J., Fang H.-L.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF467897; AAL76269.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR005054; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 206 AA; 23992 MW; 965B62F7DFFB90B9 CRC64;

Query Match
Best Local Similarity 69.7%; Score 69; DB 13; Length 206;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LEAKICHOIEYFGD 17
Db 13 LESKICHOIEYFGN 27

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:29:01; Search time 41.229 Seconds

(without alignments)
137.751 Million cell updates/sec

Title: US-09-836-073-14

Perfect score: 99
Sequence: 1 AALEKXICHQIEHYRGDF 18

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL.25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp Vertebrate.*
14: sp Unclassified.*
15: sp_rvitus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95	96.0	381	11	Q9CYB9 mus musculus
2	95	96.0	415	11	Q8BTU4 mus musculus
3	70	70.7	427	13	Q7ZTK2 xenopus lae
4	69	69.7	206	13	Q8QHS2 gallus gall
5	62	62.6	401	13	Q7ZTIO brachydanio
6	51	51.5	390	5	Q8T8V5
7	50	50.5	411	10	Q9FL36
8	50	50.5	422	10	Q9A438
9	49	49.5	381	16	Q8NYN8
10	49	49.5	381	2	P72382
11	49	49.5	391	2	P95709
12	49	49.5	391	16	Q99X57
13	49	49.5	481	10	Q8LMP9
14	48.5	48.5	928	10	Q9LJ02
15	48	48.5	91	5	O44678
16	48	48.5	545	10	O80567

17	47	47.5	119	10	Q9ZP89	Q9ZP89 neurochea
18	47	47.5	396	5	O01806	O01806 caenorhabdi
19	47	47.5	547	10	Q7XVC6	Q7XVC6 oryza sativ
20	46	46.5	568	16	Q9KJL7	Q9KJL7 vibrio chol
21	45	45.5	478	16	O87G58	O87G58 vibrio para
22	44.5	44.9	839	5	Q18841	Q18841 caenorhabdi
23	44	44.4	826	10	Q940X9	Q940X9 arabidopsis
24	44	44.4	2349	5	O81455	O81455 plasmodium
25	43.5	43.9	788	17	O58603	O58603 pyrococcus
26	43	43.4	134	16	O825D7	O825D7 streptomyce
27	43	43.4	150	10	Q9SMD8	O9SMD8 arabidopsis
28	43	43.4	239	10	Q8SMD8	O8SMD8 lammaria d
29	43	43.4	381	16	O8Z632	O8Z632 salmoneilla
30	43	43.4	386	16	O83BP5	O83BP5 coxiella bu
31	43	43.4	389	10	O8S0T8	O8S0T8 oryza sativ
32	43	43.4	523	10	Q94K80	Q94K80 arabidopsis
33	43	43.4	596	5	O8T6A8	O8T6A8 caenorhabdi
34	43	43.4	654	4	Q96NNA	Q96NNA homo sapien
35	43	43.4	654	6	O95KA5	O95KA5 macaca fasc
36	43	43.4	658	5	O18215	O18215 caenorhabdi
37	43	43.4	863	10	Q7XO50	Q7XO50 oryza sativ
38	43	43.4	1126	5	O86P80	O86P80 cryptospori
39	43	43.4	1126	5	Q7YX10	Q7YX10 homo sapien
40	43	43.4	1136	4	Q7Z6L5	Q7Z6L5 clostridium
41	42.5	42.9	577	16	Q97DB6	Q97DB6 bacteroides
42	42.5	42.9	748	16	O8AAX3	O8AAX3 brassica ol
43	42	42.4	96	10	Q9AUG1	Q9AUG1 helicobacte
44	42	42.4	200	16	Q7VIA1	Q7VIA1 lactobacilli
45	42	42.4	244	16	Q88VYAS	Q88VYAS lactobacilli

ALIGNMENTS

RESULT 1	PRELIMINARY;	PRT;	381 AA.
ID Q9CYB9	Q9CYB9		
AC Q9CYB9	Q9CYB9		
DT 01-JUN-2001	(TREMBLrel. 17, Created)		
DT 01-JUN-2001	(TREMBLrel. 17, Last sequence update)		
DT 01-OCT-2003	(TREMBLrel. 25, Last annotation update)		
DE Sjogren syndrome antigen B.			
GN SSB.			
OS Mus musculus (Mouse).			
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX NCBI_Taxid=10090;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=C57BL/6J; TISSUE=Embryo;			
RX MEDLINE=21085660; PubMed=11217551;			
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA Arakawa T., Hara A., Fukunishi Y., Komoto H., Adachi J., Fukuda S.,			
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,			
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,			
RA Kadota K., Matsumura H.A., Ashburner M., Batalov S., Casavant T.,			
RA Riechmann W., Gaasterland T., Glasi C., King B., Kochiwa H.,			
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,			
RA Schriml L.M., Stabli F., Suzuki R., Tomita M., Wagner L., Mashio T.,			
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,			
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA Lyons P., Marchionni L., Mashima U., Mazzarelli J., Mombarts P.,			
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA Sasaki K., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA Suzuki H., Toyono-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,			
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,			
RA Hayashizaki Y.,			
RT "Functional annotation of a full-length mouse cDNA collection."			
RL Nature 409:685-690(2001).			
DR EMBL; AK017822; BAB30957.1; -			
DR MGD; MGI:98423; Ssb.			

This Page Blank (uspto)

C.Species: *Vibrio cholerae*

C.Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 28-Apr-2003

C.Accession: C82379

R.Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; Chardon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Drygoli, I.; Sellers, R. L. R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A.Reference number: AB2035; M01D:20406833; PMID:10952301

A.Accession: C82379

A>Status: preliminary

A.Molecule type: DNA

A.Residues: 1-568 <HEI>

A.Cross-references: GB:AE004434; GB:AE003853; NID:g9658531; PIDN:AAF96979.1; GSPDB:GN001

A.Experimental source: serogroup O1; strain N16961; biotype El Tor

C.Genetics:

A:Gene: VCA1086

A:Map position: 2

C:Superfamily: response regulator, hnr type; response regulator homology

Query Match 46.5%; Score 46; DB 2; Length 568;

Best Local Similarity 50.0%; Pred. No. 10;

Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 LEAKICHOIEYFG 16

DB 156 LMEWCHQVEHIFG 169

RESULT 15

T20230

hypothetical protein C54G10.2 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C.Accession: T20230

R:Matthews, L.

Submitted to the EMBL Data Library, June 1996

A:Reference number: Z19240

A:Accession: T20230

A>Status: preliminary; translated from GB/EMBL/DDAU

A:Molecule type: DNA

A.Residues: 1-839 <WIL>

A.Cross-references: EMBL:Z75532; PIDN:CAA99812.1; GSPDB:GN00023; CESP:C54G10.2

A:Experimental source: clone C54G10

C:Genetics:

A:Gene: CESP:C54G10.2

A:Map position: 5

A:introns: 12/1; 34/3; 69/3; 326/1; 432/3; 535/3; 668/2

Query Match 44.9%; Score 44.5; DB 2; Length 839;

Best Local Similarity 57.9%; Pred. No. 27;

Matches 11; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

QY 3 LEAKI-----CHOIEYFG 16

DB 382 LEAKIGELSGHOIEPFG 400

Search completed: September 10, 2004, 18:02:31
Job time: 15.0782 secs

RESULT 10

T00677

hypochemical protein At2g43970 [imported] - Arabidopsis thaliana

N:Alternate names: hypochemical protein FE13.10

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 23-Mar-2001

C:Accession: T00677; G84872

R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, June 1998

A:Description: Arabidopsis thaliana chromosome II BAC FE13 genomic sequence.

A:Reference number: Z14180

A:Accession: T00677

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-529 <R0>

A:Cross-references: EMBL:AC004005; NID:G3212846; PID:G3212854

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Wolf, K.S.; Cronan, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.; Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: AB4420; MUID:20083487; PMID:10617197

A:Accession: G84872

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-529 <STO>

A:Cross-references: GB:AE002093; NID:G3212854; PID:ACC23405.1; GSPDB:GN00139

C:Genetics:

A:Gene: FE13.10; At2g43970

A:Map position: 2

A:Introns: 200/3; 228/1; 259/3; 283/3; 310/2; 344/3; 367/3; 421/3

Query Match 48.5% Score 48; DB 2; Length 529;

Best Local Similarity 66.7%; Pred. No. 4.5;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 6 KICHOIEYFGD 17

DB 196 KIVNQVEYFSD 207

RESULT 11

T30953

hypochemical protein C44E4.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000

C:Accession: T30953

R:Sammons, L.; Wohlmann, P.; Gilling, B.

submitted to the EMBL Data Library, August 1999

A:Description: The sequence of C. elegans cosmid C44E4.

A:Reference number: Z20945

A:Accession: T30953

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-396 <SAM>

A:Cross-references: EMBL:AF003140; PIDN:AAB54169.1

A:Experimental source: strain Bristol N2; clone C44E4

C:Genetics:

A:Map position: 1

A:Introns: 45/1; 114/3

A:Note: C44E4.4

Query Match 47.5% Score 47; DB 2; Length 396;

Best Local Similarity 57.1%; Pred. No. 4.9;

Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 4 EAKICHOIEYFGD 17

DB 13 DOKIKOLEYFGN 26

RESULT 12

T38937

rna binding protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T38937

R:Baccock, K.; Churruarin, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, April 1997

A:Reference number: Z21818

A:Accession: T38937

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-298 <BAD>

A:Cross-references: EMBL:Z94864; PIDN:CAB08173.1; GSPDB:GN00066; SPDB:SPAC57A10.10C

A:Experimental source: strain 972h-; cosmid c57A10

C:Genetics:

A:Gene: SPDB:SPAC57A10.10C

A:Map position: 1

A:Introns: 72/1

Query Match 46.5% Score 46; DB 2; Length 298;

Best Local Similarity 50.0%; Pred. No. 5.5;

Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 4 EAKICHOIEYFGD 17

DB 64 EAEVLKQVEYFSD 77

RESULT 13

T3542

RNA-binding protein La1 homolog - fission yeast (Schizosaccharomyces pombe)

N:Alternate names: La autoantigen; ribonucleoprotein La homolog

C:Species: Schizosaccharomyces pombe

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000

C:Accession: T43542; T43325

R:Van Horn, D.O.; Yoo, C.O.; Xue, D.; Shi, H.; Wolin, S.L.

RNA 3, 1434-1443, 1997

A:Title: The La protein in Schizosaccharomyces pombe: a conserved yet dispensable phosph

A:Reference number: Z22560; MUID:98067398; PMID:9404894

A:Accession: T43542

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-298 <VAN>

A:Cross-references: EMBL:AF022949; PIDN:AAB82145.1

R:Utsumi, R.

submitted to the EMBL Data Library, February 1998

A:Description: Screening of S. pombe cDNA library using E. coli defective in signal tran

A:Reference number: Z22428

A:Accession: T43325

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-298 <UTS>

A:Cross-references: EMBL:AB011371; PIDN:BAA24981.1

C:Genetics:

A:Gene: slal

C:Function:

A:Description: the binding of the La protein to tRNA precursors is required for the endo

C:Keywords: phosphoprotein; RNA binding

Query Match 46.5% Score 46; DB 2; Length 298;

Best Local Similarity 50.0%; Pred. No. 5.5;

Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 4 EAKICHOIEYFGD 17

DB 64 EAEVLKQVEYFSD 77

RESULT 14

G82379

response regulator VCA1086 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

R.Scherly, D.; Stutz, F.; Lin-Marq, N.; Clarkeon, S.G.

J. Mol. Biol. 231, 196-204, 1993

A>Title: La proteins from *Xenopus laevis*. cDNA cloning and developmental expression.

A:Reference number: S33817; MUID:93287095; PMID:8510143

A:Accession: S33818

A:Molecule type: mRNA

A:Residues: 1-428 <SCH>

A:Cross-references: EMBL:X68817; NID:g64873; PIDN:CA48715.1; PID:g64874

C:Comment: This protein associates with a variety of small RNA molecules, most of which

act as a transcription termination factor.

C:Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology

C:Keywords: phosphoprotein; RNA binding

F:112-178/Domain: ribonucleoprotein repeat homology <RRM>

F:113-118/Region: RNA-binding RNP motif

F:151-158/Region: RNA-binding RNP motif

F:228-428/Domain: phosphorylated #status predicted <PHY>

Query Match 66.7%; Score 66; DB 1; Length 428;

Best Local Similarity 78.6%; Pred. No. 0.0035;

Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGD 17

Db 14 DTKICQIEYFGD 27

RESULT 6

La/SS-B homolog D-la - fruit fly (*Drosophila melanogaster*)

C:Species: *Drosophila melanogaster*

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000

C:Accession: A53773

R:Bal, C.; Li, Z.; Tolias, P.P.

Mol. Cell. Biol. 14, 5123-5129, 1994

A>Title: Developmental characterization of a *Drosophila* RNA-binding protein homologous

A:Reference number: A53773; MUID:94309632; PMID:8035794

A:Accession: A53773

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-390 <BAI>

A:Cross-references: GB:U07652; NID:g464019; PIDN:AAA20518.1; PID:g464020

C:Genetics:

A:Gene: FlyBase:la

A:Cross-references: FlyBase:Fgn0011638

C:Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology

C:Keywords: leucine zipper; RNA binding

Query Match 51.5%; Score 51; DB 2; Length 390;

Best Local Similarity 64.3%; Pred. No. 1;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGD 17

Db 51 ERAITROVEYFGD 64

RESULT 7

ribonucleoprotein la - fruit fly (*Drosophila melanogaster*)

C:Species: *Drosophila melanogaster*

C>Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 20-Sep-1999

C:Accession: A53781

R:Yoo, C.J.; Wolin, S.L.

Mol. Cell. Biol. 14, 5412-5424, 1994

A>Title: La proteins from *Drosophila melanogaster* and *Saccharomyces cerevisiae*: a yeast

A:Reference number: A53781; MUID:94309661; PMID:8035818

A:Accession: A53781

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-390 <YOO>

A:Cross-references: GB:L3298; NID:g488469; PID:g488470

C:Genetics:

A:Gene: FlyBase:la

A:Cross-references: FlyBase:Fgn0011638

C:Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology

C:Keywords: RNA binding

Query Match 51.5%; Score 51; DB 2; Length 390;

Best Local Similarity 64.3%; Pred. No. 1;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGD 17

Db 51 ERAITROVEYFGD 64

RESULT 8

capsular polysaccharide synthase enzyme Caps [imported] - *Staphylococcus aureus* (strain

C:Species: *Staphylococcus aureus*

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: H89777

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc

ma, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A>Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: H89777

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-391 <KUR>

A:Cross-references: GB:BA000018; PID:g1370080; PIDN:BA841379.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: cap

C:Superfamily: lipopolysaccharide biosynthesis protein bplD

Query Match 49.5%; Score 49; DB 2; Length 391;

Best Local Similarity 72.7%; Pred. No. 2.2;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIEYFG 16

Db 366 RICEAIEYFG 376

RESULT 9

hypothetical protein C14C6.12 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T32701

R:David, M.; Wohldmann, P.; Bauer, C.; Antoniou, B.

submitted to the EMBL Data Library, December 1997

A:Description: The sequence of *C. elegans* cosmid C14C6.

A:Reference number: Z21210

A:Accession: T32701

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-166 <DAV>

A:Cross-references: EMBL:AF030051; PIDN:AB94258.1; GSPDB:GN00023; CESP:C14C6.12

A:Experimental source: strain Bristol N2; clone C14C6

C:Genetics:

A:Gene: CESP:C14C6.12

A:Map position: 5

A:Insertions: 42/1, 156/3

Query Match 48.5%; Score 48; DB 2; Length 166;

Best Local Similarity 69.2%; Pred. No. 1.4;

Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 KICHOIEYFGDF 18

Db 81 KICNVIEYMTGDF 93

A:Reference number: S03848; MUID:89202037; PMID:2468131
 A:Accession: S03848
 A:Molecule type: mRNA
 A:Residues: 1-408 <CH2>
 A:Cross-references: EMBL:X13697; NID:G36414; PIDN:CAA31985.1; PID:G36415
 R:Chambers, J.C.; Keene, J.D.
 Proc. Natl. Acad. Sci. U.S.A. 82, 2115-2119, 1985
 A:Title: Isolation and analysis of cDNA clones expressing human lupus La antigen.
 A:Reference number: A22956; MUID:85166283; PMID:3856888
 A:Accession: A22956
 A:Molecule type: mRNA
 A:Residues: 45-97, 'LK' <CH3>
 A:Cross-references: GB:J04205
 A:Note: this sequence has been revised in reference A31888
 R:Nyman, U.; Ringertz, N.R.; Petersson, I.
 Immunol. Lett. 22, 65-72, 1989
 A:Title: Demonstration of an amino terminal La epitope recognized by human anti-La sera.
 A:Reference number: A61051; MUID:89379261; PMID:2476379
 A:Accession: A61051
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-19, 'E', '21-47' <NTM>
 R:Sturges, A.D.; Peterson, M.G.; McNeillage, L.J.; Whittingham, S.; Coppel, R.L.
 J. Immunol. 140, 3212-3218, 1988
 A:Title: Characteristics and epitope mapping of a cloned human autoantigen La.
 A:Reference number: S11013; MUID:88199081; PMID:2452201
 A:Accession: S11013
 A:Molecule type: mRNA
 A:Residues: 'E', '55-287', 'V', '289-408' <STU>
 A:Cross-references: EMBL:M20328; NID:G337456; PIDN:AAA36577.1; PID:G337457
 R:Kohsaka, H.; Yamamoto, K.; Fujii, H.; Mura, H.; Miyasaka, N.; Nishio, K.; Miyamoto, J.
 Clin. Invest. 85, 1566-1574, 1990
 A:Title: Fine epitope mapping the human SS-B/La protein: Identification of a distinct autoantigen.
 A:Reference number: 15553; MUID:90237237; PMID:1692037
 A:Accession: 15553
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 81-107 <RES>
 A:Cross-references: GB:M35261; NID:G338491; PIDN:AAA36652.1; PID:G338495
 A:Accession: 170205
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 174-224 <RE2>
 A:Cross-references: GB:M35263; NID:G338492; PIDN:AAA36653.1; PID:G338496
 A:Accession: 170206
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 279-342 <RE3>
 A:Cross-references: GB:M35262; NID:G338493; PIDN:AAA36654.1; PID:G338497
 C:Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.
 C:Genetics:
 A:Gene: GDB:SSB
 A:Cross-references: GDB:125359; OMIM:109090
 A:Map position: 2
 A:Introns: 22/3; 57/2; 115/3; 151/3; 185/2; 209/2; 223/3; 264/3; 380/2
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:112-118/Domain: ribonucleoprotein repeat homology <RMM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:228-408/Domain: phosphorylated #status experimental <PHY>

Query Match 100.0%; Score 99; DB 1; Length 408;
 Best Local Similarity 100.0%; Pred. No. 9.8e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AALEAKICHOIEYFGDF 18
 |||||
 Db 11 AALEAKICHOIEYFGDF 28

JC1494
 ribonucleoprotein La - rat
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 24-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: JC1494; S25145
 R:Semsel, I.; Troester, H.; Bartsch, H.; Schwemle, M.; Igloi, G.L.; Bachmann, M.
 Gene 126, 265-268, 1993
 A:Title: Isolation of rat cDNA clones coding for the autoantigen SS-B/La: Detection of B
 A:Reference number: JC1494; MUID:93246255; PMID:7916708
 A:Accession: JC1494
 A:Molecule type: mRNA
 A:Residues: 1-415 <SEM>
 A:Cross-references: GB:X67859; NID:G55778; PIDN:CAA48043.1; PID:G55779
 A:Experimental source: liver
 C:Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:112-118/Domain: ribonucleoprotein repeat homology <RMM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:227-415/Domain: phosphorylated #status predicted <PHY>

Query Match 100.0%; Score 99; DB 1; Length 415;
 Best Local Similarity 100.0%; Pred. No. 1e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AALEAKICHOIEYFGDF 18
 |||||
 Db 11 AALEAKICHOIEYFGDF 28

RESULT 4

S33817
 ribonucleoprotein La.B - African clawed frog
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33817; S28544
 R:Scherer, D.; Stutz, F.; Lun-Marg, N.; Clarkson, S.G.
 J. Mol. Biol. 231, 196-204, 1993
 A:Title: La protein from Xenopus laevis. cDNA cloning and developmental expression.
 A:Reference number: S33817; MUID:93287095; PMID:8510143
 A:Accession: S33817
 A:Molecule type: mRNA
 A:Residues: 1-427 <SCH>
 A:Cross-references: EMBL:X68818; NID:G64875; PIDN:CAA48716.1; PID:G64876
 C:Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:112-117/Domain: ribonucleoprotein repeat homology <RMM>
 F:113-117/Region: RNA-binding RNP2 motif
 F:150-157/Region: RNA-binding RNP1 motif
 F:227-427/Domain: phosphorylated #status predicted <PHY>

Query Match 70.7%; Score 70; DB 1; Length 427;
 Best Local Similarity 80.0%; Pred. No. 0.00074;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 LEAKICHOIEYFGD 17
 |||||
 Db 12 LDRICHOIEYFGD 26

RESULT 5

S33818
 ribonucleoprotein La.A - African clawed frog
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33818; S28545

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:33:11; Search time 14.0782 Seconds
(without alignments)
122.988 Million cell updates/sec

Title: US-09-836-073-14

Perfect score: 99

Sequence: 1 AALEAKICHQIEYFGDF 18

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

283366

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	100.0	404	1 S03849	ribonucleoprotein
2	99	100.0	408	1 A31888	ribonucleoprotein
3	99	100.0	415	1 J01494	ribonucleoprotein
4	70	70.7	427	1 S33817	ribonucleoprotein
5	66	66.7	428	1 S33818	ribonucleoprotein
6	51	51.5	390	2 A53773	la/SS-B homolog D-
7	51	51.5	390	2 A53781	ribonucleoprotein
8	49	49.5	391	2 H89777	capsular polysacch
9	48	48.5	321	2 T32701	hypothetical prote
10	48	48.5	529	2 T00677	hypothetical prote
11	47	47.5	396	2 T30953	hypothetical prote
12	46	46.5	298	2 T38937	rna binding protei
13	46	46.5	298	2 T43542	RNA-binding protei
14	46	46.5	568	2 C82379	response regulator
15	44	44.9	839	2 T20230	hypothetical prote
16	43	43.9	788	2 A71076	hypothetical prote
17	43	43.4	150	2 A64488	hypothetical prote
18	43	43.4	381	2 AB0734	probable bacterio
19	43	43.4	506	2 T14907	trans-cinnamate 4-
20	43	43.4	658	2 T19487	hypothetical prote
21	42	42.9	577	2 D97337	mismatch repair pr
22	42	42.4	242	2 A42478	glutamine transpor
23	42	42.4	422	2 D72313	hypothetical prote
24	42	42.4	424	2 T46197	hypothetical prote
25	42	42.4	482	2 B35843	lipopolysaccharide
26	42	42.4	505	2 B90181	Na+/H+ antiporter
27	42	42.4	541	2 H71887	hypothetical prote
28	42	42.4	542	2 G64637	hypothetical prote
29	42	42.4	658	2 D96656	hypothetical prote

30	42	42.4	1131	2 T15617	hypothetical prote
31	41	41.4	87	2 C84494	hypothetical prote
32	41	41.4	251	2 B40969	folate-binding pro
33	41	41.4	419	2 AH0417	integrinase [impor
34	41	41.4	469	2 C70357	hypothetical prote
35	41	41.4	506	2 F85016	probable RING zinc
36	41	41.4	573	1 CSBT	catalase (EC 1.11.
37	41	41.4	1156	2 T37411	RNA polymerase sub
38	41	41.4	1164	1 RNVT8T	DNA-directed RNA p
39	41	41.4	1164	1 RNVT8T	DNA-directed RNA p
40	41	41.4	1164	2 T28565	DNA-directed RNA p
41	41	41.4	1164	2 F72166	A25R protein - var
42	41	41.4	1164	2 G36850	DNA-directed DNA p
43	41	41.4	1220	2 DJBEC3	DNA-directed DNA p
44	41	41.4	1220	2 T42573	DNA-directed DNA p
45	41	41.4	1221	1 HJNVAV	helicase (EC 3.6.1

ALIGNMENTS

RESULT 1

S03849
ribonucleoprotein la - bovine
N/Alternate names: autoantigen SS-B/la, ribonucleoprotein SS-B
C/Species: Bos primigenius taurus (cattle)
C/Date: 31-Dec-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C/Accession: S03849
R/Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.
Nucleic Acids Res. 17, 2233-2244, 1989
A/Title: Ribonucleoprotein SS-B/la belongs to a protein family with consensus sequences
A/Reference number: S03848; PMID:89202037; PMID:2468131
A/Accession: S03849
A/Molecule type: mRNA
A/Residues: 1-404 <CHA>
A/Cross-References: EMBL:X13698; NID:g755; PIDN:CAJ1986.1; PID:g756
A/Note: Part of this sequence was confirmed by protein sequencing
C/Comment: This protein associates with a variety of small RNA molecules, most of which
ay act as a transcription termination factor.
C/Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology
C/Keywords: blocked amino end, phosphoprotein; RNA binding
F/117-178/Domain: ribonucleoprotein repeat homology <RRM>
F/113-118/Region: RNA-binding RNP2 motif
F/151-158/Region: RNA-binding RNP1 motif
F/228-404/Domain: phosphorylated #status predicted <PHY>

Query Match 100.0%; Score 99; DB 1; Length 404;

Best Local Similarity 100.0%; Pred. No. 9.7e-09; Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYFGDF 18
DB 11 AALEAKICHQIEYFGDF 28

RESULT 2

A31888
ribonucleoprotein la - human
N/Alternate names: autoantigen SS-B/la; ribonucleoprotein SS-B; Sjogren syndrome antigen
C/Species: Homo sapiens (man)
C/Date: 21-May-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C/Accession: A31888; S03848; A22956; A61051; S11013; I55553; I70205; I70206; A31273
R/Chambers, J.C.; Kenan, D.; Martin, B.J.; Keene, J.D.
J. Biol. Chem. 263, 18043-18051, 1988
A/Title: Genomic structure and amino acid sequence domains of the human la autoantigen.
A/Reference number: A31888; PMID:89053970; PMID:3192525
A/Accession: A31888
A/Molecule type: mRNA
A/Residues: 1-408 <CHA>
A/Cross-References: GB:04205; NID:g178686; PIDN:AAA51885.1; PID:g178687
R/Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.
Nucleic Acids Res. 17, 2233-2244, 1989
A/Title: Ribonucleoprotein SS-B/la belongs to a protein family with consensus sequences

DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La.dom.
 DR InterPro; IPR00504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; trm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 SQ SEQUENCE 367 AA; 40389 MW; 129BF02581DD1148 CRC64;

Query Match 46.4%; Score 45; DB 10; Length 367;
 Best Local Similarity 60.0%; Pred. No. 23;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Oy 3 LEAKICHOIEYOFGD 17
 |||||
 Db 147 LRLKIVKQVEYOFD 161

RESULT 15

O9LHL3 PRELIMINARY; PRT; 455 AA.
 AC O9LHL3;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE RNA-binding protein-like.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxId=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu B., Tabata S.;
 RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RX MEDLINE=20363099; PubMed=10907853;
 RA Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
 RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,
 RT TAC and BAC clones."
 RL DNA Res. 7:217-221(2000).
 DR EMBL; AP002041; BAB02607.1; -;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La.dom.
 DR InterPro; IPR00504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; trm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 SQ SEQUENCE 455 AA; 48958 MW; E88117B2C33BA9 CRC64;

Query Match 46.4%; Score 45; DB 10; Length 455;
 Best Local Similarity 60.0%; Pred. No. 28;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Oy 3 LEAKICHOIEYOFGD 17
 |||||
 Db 144 LRLKIVKQVEYOFD 158

Search completed: September 10, 2004, 18:00:06
 Job time : 44.229 secs

DR SMART: SM00360; RM; 1.
DR PROSITE: PS50102; RM; 1.
SQ SEQUENCE 343 AA; 37628 MW; 655FCE0879FBBD7B CRC64;

Query Match
Best Local Similarity 47.4%; Score 46; DB 10; Length 343;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 AALBAKICHOIEYQFD 17
DB 111 ADIAQKIKQVEYQFSD 127

RESULT 12

ID Q8U3P2 PRELIMINARY; PRT; 395 AA.
AC Q8U3P2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dehydrogenase.
GN PF0415.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_Taxid=2261;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
EMBL: AEO10164; AAL80539.1; -
DR GO: 0006118; P:electron transport; IEA.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR SMART: SM00564; PQQ; 5.
KW Complete proteome.
SQ SEQUENCE 395 AA; 43482 MW; 946EC9C97EE327D4 CRC64;

Query Match
Best Local Similarity 47.4%; Score 46; DB 17; Length 395;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ALEAKICHOIEYQ 14
DB 19 AWEKICENIEYQ 31

RESULT 13

ID Q9VU19 PRELIMINARY; PRT; 519 AA.
AC Q9VU19;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CG10006 protein.
GN CG10006.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananthakrishnan P.G., Scherer S.E., Li P.W., Hoskins R.A., Gill R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Anil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck P., Brocktein P., Broctier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison A.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Switzkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AEO03532; AAF49687.1; -
DR FlyBase: FBgn0036461; CG10006.
DR GO: 0016020; C:membrane; IEA.
DR GO: 0046873; F:metal ion transporter activity; IEA.
DR GO: 0030001; P:metal ion transport; IEA.
DR InterPro: IPR003689; Zn_transpl_zfp.
DR Pfam: PF02535; Zfp; 1.
SQ SEQUENCE 519 AA; 56866 MW; 670677CC0E125DAC CRC64;

Query Match
Best Local Similarity 47.4%; Score 46; DB 5; Length 519;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 5 AKICHOIEYQFGDF 18
DB 395 AVLCHELPHELDGF 408

RESULT 14

ID Q9SEU9 PRELIMINARY; PRT; 367 AA.
AC Q9SEU9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RNA-binding protein homolog.
GN RBP1.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_Taxid=3708;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=cv. topas; TISSUE=Late-uninucleate microspores;
RA Smekal P., Janotov I., Pechan P.M.;
RT "A novel Brassica napus L. pollen-specific gene belongs to a nucleic-acid-binding protein family."
RL Sex. Plant Reprod. 13:127-134(2000).
DR EMBL: AF094825; AAF00075.2; -
DR GO: 0005634; C:nucleus; IEA.
DR GO: 0003723; F:RNA binding; IEA.

```

RA Sugiyama T., Irie R., Otsuki T., Sato H., Oca T., Wakamatsu A.,
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
RA Kikuchi H., Kanda K., Magatsuna M., Murakawa K., Kanehori K.,
RA Takahashi-Fuji A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahara K., Maehuo Y., Nagai K., Isogai T.,
RT "NEBO human cDNA sequencing project."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK055061; BAB70848.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0046873; F:metal ion transporter activity; IEA.
DR GO: GO:0030001; P:metal ion transport; IEA.
DR InterPro: IPR003689; Zn_transp_zip.
DR Pfam: PF02535; Zip; 1.
KW Hypothetical protein.
SQ SEQUENCE 654 AA; 72805 MW; D5E1AD30707C88F7 CRC64;

Query Match
Best Local Similarity 48.5%; Score 47; DB 4; Length 654;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Oy 5 AKICHOIEYQFGDF 18
Db 539 AILCHEIPHEMGDF 552

RESULT 9
O9SKAS PRELIMINARY; PRT; 654 AA.
AC O9SKAS;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
OS Hypothetical protein.
DS Macaca faecicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Medulla oblongata;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB063029; BAB60771.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0046873; F:metal ion transporter activity; IEA.
DR GO: GO:0030001; P:metal ion transport; IEA.
DR InterPro: IPR003689; Zn_transp_zip.
DR Pfam: PF02535; Zip; 1.
KW Hypothetical protein.
SQ SEQUENCE 654 AA; 72793 MW; 99E4D4893FFB4CBC CRC64;

Query Match
Best Local Similarity 48.5%; Score 47; DB 6; Length 654;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Oy 5 AKICHOIEYQFGDF 18
Db 539 AILCHEIPHEMGDF 552

RESULT 10
O8XW47 PRELIMINARY; PRT; 169 AA.
AC O8XW47;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Probable 2-amino-4-hydroxy-6-hydroxymethylidihydropteridine

```

```

DE pyrophosphokinase, 8-dihydro-6-hydroxymethylpterin-pyrophosphokinase
DE protein (EC 2.7.6.3).
GN FOLK OR RSC2628 OR RS00932.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gauzy J., Mangelot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Chisme N., Claudel-Renard C., Cunnc S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiek T.,
RA Siglier P., Thebaud P., Whalen W., Winkler P., Levy M.,
RA Weissbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
RL Nature 415:497-502(2002).
DR EMBL: AL646071; CAD16335.1; -.
DR GO: GO:0003848; F:2-amino-4-hydroxy-6-hydroxymethylidihydropte. .; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0009396; P:folic acid and derivative biosynthesis; IEA.
DR InterPro: IPR000550; Hppk.
DR Pfam: PF01288; HPPK; 1.
DR TIGRfam: TIGR01498; folk; 1.
DR PROSITE: PS00794; HPPK; 1.
KW transferase; Complete proteome.
SQ SEQUENCE 169 AA; 18179 MW; 829CD87425C21F36 CRC64;

Query Match
Best Local Similarity 47.4%; Score 46; DB 16; Length 169;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 6 KICHOIEYQFG 16
Db 71 RICHIEDQFG 81

RESULT 11
O94LD0 PRELIMINARY; PRT; 343 AA.
AC O94LD0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative RNA-binding protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriatroidae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nidponbare;
RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,
RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsiurin T.,
RA Riggs F., Hsiao J., Zismann V., Blunt S., Pai G., Vanaken S.E.,
RA Utechtack T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,
RA White O., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNbA0026A15 genomic sequence."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC064404; AAK50577.1; -.
DR Gramene; O94LD0; -.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR InterPro: IPR002344; Lupus_La.
DR InterPro: IPR006630; Lupus_La_dom.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF05383; La; 1.
DR Pfam: PF00076; xrm; 1.
DR PRINTS: PR00302; LUPUSLA.
DR SMART: SMO0715; LA; 1.

```

RESULT 5

ID Q7ZT10 PRELIMINARY; PRT; 401 AA.

AC Q7ZT10;

DT 01-JUN-2003 (TEMBLrel. 24, Created)

DT 01-JUN-2003 (TEMBLrel. 25, Last sequence update)

DE 01-OCT-2003 (TEMBLrel. 25, Last annotation update)

DE Similar to sjogren syndrome antigen B (Autoantigen La).

OS Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

OC NCBI_TaxId=7955;

OX NCBI_TaxId=7955;

RN (1)

RP SEQUENCE FROM N.A.

RC TISSUE=Body;

RA Strausberg R.;

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

ID EMBL; BC045392; AAH45392.1; -.

DR GO; GO:0005634; C:nucleus; IEA.

DR GO; GO:0003723; F:RNA binding; IEA.

DR InterPro; IPR002344; Iupus La.

DR InterPro; IPR006630; Iupus La dom.

DR InterPro; IPR000504; RNA_rec_mot.

DR Pfam; PF00076; rrm; 1.

DR PRINTS; PR00302; LUPUSLA.

DR SMART; SM00715; LA; 1.

DR SMART; SM00360; RRM; 1.

DR PROSITE; PS0102; RRM; 1.

DR PROSITE; PS00030; RRM_RNP_1; 1.

DR PROSITE; PS00030; RRM_RNP_1; 1.

SQ SEQUENCE 401 AA; 46138 MW; 2EF032FDD316291 CRC64;

Query Match 55.7%; Score 54; DB 13; Length 401;

Best Local Similarity 66.7%; Pred. No. 0.71;

Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 LEAKICHOIEYQFGD 17

DB 10 LEKVAEQIEYVFGD 24

RESULT 6

ID Q44678 PRELIMINARY; PRT; 91 AA.

AC Q44678;

DT 01-JUN-1998 (TEMBLrel. 06, Created)

DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)

DE Hypothetical protein C14C6.12.

GN C14C6.12.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea; Rhabditidae; Pelodierinae; Caenorhabditis.

OC NCBI_TaxId=6239;

OX NCBI_TaxId=6239;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RX MEDLINE=99069613; PubMed=9851916;

RA Wilson R.;

RT "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium.";

RL Science 282:2012-2018 (1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA David W., Wohlmann P., Bauer C., Antoniou B.;

RT "The sequence of C. elegans cosmid C14C6.";

RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Waterston R.;

RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Wilson R.;

RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF039051; AAB94258.2; -.

KW Hypothetical protein.

SQ SEQUENCE 91 AA; 10407 MW; D16D2BD4011F8762 CRC64;

Query Match 50.5%; Score 49; DB 5; Length 91;

Best Local Similarity 69.2%; Pred. No. 1.1;

Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 KICHOIEYQFGDF 18

DB 6 KICNVIEYVTDGF 18

RESULT 7

ID Q7XT99 PRELIMINARY; PRT; 1560 AA.

AC Q7XT99;

DT 01-OCT-2003 (TEMBLrel. 25, Created)

DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)

DE OSJNB0008A08.11 protein.

GN OSJNB0008A08.11.

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Eriarbioideae; Oryzaceae; Oryza.

OC NCBI_TaxId=4530;

RN (1)

RP SEQUENCE FROM N.A.

RC Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Zhao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H., Gu J.L., Chen S.T., Ni L., Zhu F.H., Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X., Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Wang Q.J., Zhang L., Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H., Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Hong G.F.;

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

RX EMBL; AL060589; CAB01603.1; -.

SQ SEQUENCE 1560 AA; 177753 MW; 172320FF3797C6F0 CRC64;

Query Match 49.5%; Score 48; DB 10; Length 1560;

Best Local Similarity 46.7%; Pred. No. 31;

Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALEAKICHOIEYQFG 16

DB 1136 AVEFQWCHRYMWFQFG 1150

RESULT 8

ID Q96NN4 PRELIMINARY; PRT; 654 AA.

AC Q96NN4;

DT 01-DEC-2001 (TEMBLrel. 19, Created)

DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)

DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)

DE Hypothetical protein FLJ30499.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OC NCBI_TaxId=9606;

OX NCBI_TaxId=9606;

RN (1)

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H.,

DR GO:0005634; C:nucleus; IDA.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dom.
 DR InterPro; IPR005054; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 SQ SEQUENCE 381 AA; 43891 MW; 2E2DEF1452C0F0E9 CRC64;

Query Match 89.7%; Score 87; DB 11; Length 381;
 Best Local Similarity 94.1%; Pred. No. 1.4e-06;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 ALEAKICHOIEYQFDF 18
 |||||
 Db 12 ALEAKICHOIEYFGDF 28

RESULT 2

O8BTU4 PRELIMINARY; PRT; 415 AA.
 AC O8BTU4;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Sjogren syndrome antigen B.
 GN SSB.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NO; TISSUE=Thymus;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 DR EMBL; AK088677; BAC40498.1; -.
 DR MGD; MGI:98423; Seb.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dom.
 DR InterPro; IPR005054; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 SQ SEQUENCE 415 AA; 47657 MW; A7545C7686AC6363 CRC64;

Query Match 89.7%; Score 87; DB 11; Length 415;
 Best Local Similarity 94.1%; Pred. No. 1.6e-06;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 ALEAKICHOIEYQFDF 18
 |||||
 Db 12 ALEAKICHOIEYFGDF 28

RESULT 3

O7ZTK2 PRELIMINARY; PRT; 427 AA.
 AC O7ZTK2;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Similar to lupus LA protein homolog B.

OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Strausberg R.,
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC046654; AAH4654.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dom.
 DR InterPro; IPR005054; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 2.
 DR PROSITE; PS50102; RRM; 1.
 SQ SEQUENCE 427 AA; 48996 MW; 1E7CD82D8AB9C69A CRC64;

Query Match 63.9%; Score 62; DB 13; Length 427;
 Best Local Similarity 73.3%; Pred. No. 0.032;
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 3 LEAKICHOIEYQFDF 17
 |||||
 Db 12 LDTKICHOIEYFGDF 26

RESULT 4

O8OH15 PRELIMINARY; PRT; 206 AA.
 AC O8OH15;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Untranslated region binding-protein.
 GN UBP.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA L'Euyer T.J., Fang H.-L.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF467897; AAL76269.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dom.
 DR InterPro; IPR005054; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 SQ SEQUENCE 206 AA; 23992 MW; 965B62F7DFB90E9 CRC64;

Query Match 62.9%; Score 61; DB 13; Length 206;
 Best Local Similarity 73.3%; Pred. No. 0.022;
 Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 3 LEAKICHOIEYQFDF 17
 |||||
 Db 13 LESKICHOIEYFGDF 27

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:29:01; Search time 41.229 Seconds
(without alignments)
137.751 Million cell updates/sec

Title: US-09-836-073-12

Perfect score: 97

Sequence: 1 AALEAKICHQIEYQFGDF 18

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL.25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp Vertebrate:*
- 14: sp Unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	89.7	381	11	Q9CYB9 mus musculus
2	87	89.7	415	11	Q8BTU4 mus musculus
3	62	63.9	427	13	Q7ZTK2 xenopus lae
4	61	62.9	206	13	Q8QHI5 gallus gall
5	54	55.7	401	13	Q7ZTI0 brachydanio
6	49	50.5	91	5	Q44678 caenorhabdi
7	48	49.5	1560	10	Q7X199 oryza sativ
8	47	48.5	654	6	Q96NNA homo sapien
9	47	48.5	654	6	Q9SKR5 macaca fasc
10	46	47.4	169	16	Q8XK47 ralestonia s
11	46	47.4	343	10	Q941D0 oryza sativ
12	46	47.4	395	17	Q8UJ32 pyrococcus
13	46	47.4	519	5	Q9VUI9 dirosophila
14	45	46.4	367	10	Q9SEU9 brassica na
15	45	46.4	455	10	Q9JHL3 arabidopsis
16	45	46.4	1456	10	Q9PFN9 oryza sativ

17	45	46.4	1456	10	Q7X975 oryza sativ
18	44	45.4	96	10	Q9AUG1
19	44	45.4	129	4	Q9EM13 homo sapien
20	44	45.4	149	4	Q9NX22 homo sapien
21	44	45.4	160	4	Q8NC35 homo sapien
22	44	45.4	228	11	Q8C0L2 mus musculus
23	44	45.4	230	11	Q8CHL4 mus musculus
24	44	45.4	381	11	Q8BX42 mus musculus
25	44	45.4	382	4	Q96HP5 homo sapien
26	44	45.4	433	4	Q8IXR3 homo sapien
27	44	45.4	505	11	Q8R518 mus musculus
28	44	45.4	535	11	Q9D856 mus musculus
29	44	45.4	535	11	Q9D909 mus musculus
30	44	45.4	539	4	Q8N6Y3 homo sapien
31	44	45.4	567	10	Q7X6P9 oryza sativ
32	44	45.4	568	16	Q9KX17 vibrio chol
33	44	45.4	571	5	Q7Z1R8 plasmodium
34	44	45.4	582	11	Q7T0E0 mus musculus
35	44	45.4	626	4	Q9NXC4 homo sapien
36	44	45.4	647	4	Q9H6T8 homo sapien
37	44	45.4	660	11	Q9DAT9 mus musculus
38	44	45.4	661	11	Q8OTG2 mus musculus
39	44	45.4	701	5	Q9VSL7 dirosophila
40	44	45.4	749	4	Q13433 homo sapien
41	44	45.4	765	11	Q8C145 mus musculus
42	44	45.4	765	11	Q7TPP9 mus musculus
43	44	45.4	835	4	Q9JLF5 homo sapien
44	44	45.4	1230	10	Q9JLD6 oryza sativ
45	44	45.4	1381	10	Q84J29 oryza sativ

ALIGNMENTS

RESULT 1	ID	Q9CYB9	PRELIMINARY	PRT	381 AA.
AC	Q9CYB9	01-JUN-2001 (TREMBL)	17, Created		
DT	01-JUN-2001 (TREMBL)	17, Last sequence update			
DE	01-OCT-2003 (TREMBL)	25, Last annotation update			
DE	Sjogren syndrome antigen B.				
GN	SSB.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6J; TISSUE=Embryo;				
RX	MEDLINE=21085660; PubMed=11217851;				
RA	Kawai U., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,				
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,				
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamahata I.,				
RA	Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,				
RA	Kadora K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,				
RA	Pfeichmann W., Gaasterland T., Giusti C., King B., Kochiwa H.,				
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,				
RA	Schirml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Mashio T.,				
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Batsch G.,				
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,				
RA	Trombstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,				
RA	Justincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,				
RA	Lyons P., Marchionni L., Mashima U., Mazzerelli U., Mombauts P.,				
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,				
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,				
RA	Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitlaker C., Wilming L.,				
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,				
RA	Hayashizaki Y.,				
RT	"Functional annotation of a full-length mouse cDNA collection."				
RL	Nature 409:85-90(2001).				
DR	EMBL; AK017822; BAB30957.1; -				
DR	MGI; MGI:98423; Ssb.				

RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
LT2.";
RU Nature 413:852-856(2001).
CC -1- FUNCTION: Mediates magnesium influx to the cytosol.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + Mg(2+) (Out) = ADP + phosphate +
CC Mg(2+) (In).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the cation transport ATPases family (P-type
CC ATPases). Subfamily IIB.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U07843; AAA68988.1; -.
CC EMBL; AE008909; AAL23275.1; -.
CC PIR; B57147; B57147.
CC StryGene; SG10230; mgTA.
DR InterPro; IPR006415; ATPase-IIB_Mg.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR004014; Cation_ATPase_N.
DR InterPro; IPR008250; E1-E2_ATPase_Reg.
DR InterPro; IPR000695; H_ATPase.
DR Pfam; PF00690; Cation_ATPase_N; 1.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolyase; 1.
DR PRINTS; PR00119; CRYPTASE.
DR PRINTS; PR00120; HATPASE.
DR TIGRFAMs; TIGR01524; ATPase-IIB_Mg; 1.
DR TIGRFAMs; TIGR01494; ATPase_P-type; 5.
DR PROSITE; PS00154; ATPase_E1-E2; 1.
KW Hydrolyase; Magnesium; Transmembrane; Phosphorylation; ATP-binding;
KW Complete proteome.
KV
FT DOMAIN 1 98
FT TRANSMEM 1 119
FT DOMAIN 120 120
FT TRANSMEM 121 141
FT DOMAIN 142 291
FT TRANSMEM 292 312
FT DOMAIN 313 321
FT TRANSMEM 322 339
FT DOMAIN 340 699
FT TRANSMEM 700 719
FT DOMAIN 720 728
FT TRANSMEM 729 748
FT DOMAIN 749 770
FT TRANSMEM 771 794
FT DOMAIN 795 803
FT TRANSMEM 804 822
FT DOMAIN 823 835
FT TRANSMEM 836 855
FT DOMAIN 856 870
FT TRANSMEM 871 890
FT DOMAIN 891 902
FT MOD_RES 379 379
FT METAL 335 335
FT METAL 645 645
FT METAL 649 649
FT METAL 713 713
FT METAL 738 738
FT METAL 742 742
SQ SEQUENCE 902 AA; 99782 MW; 3D2712E9A074C957 CRC64;

Query Match 42.3%; Score 41; DB 1; Length 902;
Best Local Similarity 46.2%; Pred. No. 37;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 AKVCHQIEVGFCD 17

DB 581 AKVCHQIEVGLDAGD 593

Search completed: September 10, 2004, 17:53:05
Job time : 8.24022 secs

RA Satou M., Tamse R., Vayenberg M., Wallender E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Empirical analysis of transcriptional activity in the Arabidopsis
 RT genome.";
 RL Science 302:842-846(2003).
 CC -1- FUNCTION: Essential for proper development of leaves and floral
 CC organs, and formation of axillary meristems.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=1;
 CC Comment=Number of isoforms are produced. According to EST
 CC sequences;
 CC Name=1;
 CC isoId=O04379-1; Sequence=Displayed;
 CC -1- TISSUE SPECIFICITY: Widely expressed at low levels.
 CC -1- DEVELOPMENTAL STAGE: Expressed throughout all developmental
 CC stages.
 CC -1- SIMILARITY: Belongs to the argonaute family.
 CC -1- SIMILARITY: Contains 1 PAZ domain.
 CC -1- SIMILARITY: Contains 1 Piwi domain.
 CC -1- CAUTION: Ref.2 (AAF79718) sequence differs from that shown due to
 CC erroneous gene model prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL, U91995; AAC18440.1; -;
 CC DR EMBL, AC007932; AAD49755.1; -;
 CC DR EMBL, AC020889; AAF79718.1; ALT_SEQ.
 CC DR EMBL, E7000941; AAN41341.1; -;
 CC DR InterPro: IPR003100; PAZ.
 CC DR InterPro: IPR003165; Piwi.
 CC DR Pfam: PF02170; PAZ; 1.
 CC DR Pfam: PF02171; Piwi; 1.
 CC DR PROSITE: PS50821; PAZ; 1.
 CC DR PROSITE: PS50822; Piwi; 1.
 CC KM Developmental protein: Alternative splicing.
 CC FT DOMAIN 391 501 PAZ.
 CC FT DOMAIN 676 997 Piwi.
 CC FT DOMAIN 13 104 GLY-RICH.
 CC SQ SEQUENCE 1048 AA; 116190 MW; 3B5146343A09C541 CRC64;
 CC -----
 CC Query Match 43.3%; Score 42; DB 1; Length 1048;
 CC Best Local Similarity 46.2%; Pred. No. 29;
 CC Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 CC -----
 CC QY 3 LEAKICHOIEYQF 15
 CC Db 915 VDSKICHPTEFDF 927
 CC -----
 CC RESULT 14
 CC TCNO PETCR STANDARD; PRT; 506 AA.
 CC AC 043033;
 CC DT 15-DEC-1998 (Rel. 37, Created)
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
 CC DE Trans-cinnamate 4-monoxygenase (EC 1.14.13.11) (Cinnamic acid
 CC 4-hydroxylase) (CA4H) (C4H) (P450CA4) (Cytochrome P450 73).
 CC GN CYP7A10 OR CYP73.
 CC OS Petroselinum crispum (Parsley) (Petroselinum hortense).
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; asterids;
 CC Campanulids; Apiales; Apiaceae; Apioideae; apioid superclade;
 CC OC Apium clade; Petroselinum.
 CC OX NCBI_TaxID=4043;
 CC [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=95320184; PubMed=7597051;
 RA Logemann E., Pariske M., Hahlbrock K.;
 RT "Modes of expression and common structural features of the complete
 RT phenylalanine ammonia-lyase gene family in parsley.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:5905-5909(1995).
 CC -1- FUNCTION: Controls carbon flux to pigments essential for
 CC pollination or UV protection, to numerous phytoalexins synthesized
 CC by plants when challenged by pathogens, and to lignins.
 CC -1- CATALYTIC ACTIVITY: Trans-cinnamate + NADPH + O(2) = 4-
 CC hydroxycinnamate + NADP(+) + H(2)O.
 CC -1- PATHWAY: Lignin biosynthesis.
 CC -1- PATHWAY: Phenylpropanoid metabolism; second step.
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL, L38898; AAC1660.1; -;
 CC DR PIR, T14907; T14907.
 CC DR InterPro: IPR001128; Cytochrome_P450.
 CC DR Pfam: PR00067; P450; 1.
 CC DR PRINTS; PR00385; P450.
 CC DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 CC KM Oxidoreductase; Monooxygenase; Heme; NADP.
 CC FT METAL 448 448 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 CC SQ SEQUENCE 506 AA; 58047 MW; 32F0EE959D69CCF CRC64;
 CC -----
 CC Query Match 42.3%; Score 41; DB 1; Length 506;
 CC Best Local Similarity 42.9%; Pred. No. 20;
 CC Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 CC -----
 CC QY 5 AKICHOIEYQFDF 18
 CC Db 215 SRLAQSEYHFGDF 228
 CC -----
 CC RESULT 15
 CC ATMA_SALTY STANDARD; PRT; 902 AA.
 CC AC P36640;
 CC DT 01-JUN-1994 (Rel. 29, Created)
 CC DT 01-JUN-1994 (Rel. 29, Last sequence update)
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
 CC DE Mg(2+) transport ATPase, P-type 1 (EC 3.6.3.2).
 CC GN MGT4 OR STM4456.
 CC OS Salmonella typhimurium.
 CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC OC Enterobacteriaceae; Salmonella.
 CC OX NCBI_TaxID=602;
 CC [1]
 CC RN SEQUENCE FROM N.A.
 CC RP STRAIN=LT2;
 CC RC MEDLINE=95270580; PubMed=7751273;
 CC RA Tao T., Shavelly M.D., Farr S.G., Maguire M.E.;
 CC "Magnesium transport in Salmonella typhimurium: mgtA encodes a P-type
 CC ATPase and is regulated by Mg2+ in a manner similar to that of the
 CC mgtB P-type ATPase.";
 CC RT mgtB P-type ATPase.";
 CC RL J. Bacteriol. 177:2654-2662(1995).
 CC [2]
 CC RN SEQUENCE FROM N.A.
 CC RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
 CC MEDLINE=21534948; PubMed=11677609;
 CC RA McCelland W., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 CC Courtney L., Potwilk S., Ali J., Dante M., Du F., Hou S., Layman D.,
 CC Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 CC Ryan E., Sun H., Flora L., Miller W., Stoneking T., Nhan M.,
 CC Waterston R., Wilson R.K.;

OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OK NCBI_TaxID=4932;
 RN [1]
 RN SEQUENCE OF 12-573 FROM N.A.
 RX MEDLINE=87053966; PubMed=3536508;
 RA Hartig A., Ruis H.;
 RT "Nucleotide sequence of the Saccharomyces cerevisiae CTT1 gene and
 RT deduced amino-acid sequence of yeast catalase T.";
 RL Eur. J. Biochem. 160:487-490 (1986).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP Medler H., Scharfe M., Medler E., Manburt R.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RA Hernandez K., Weber N., Widfl P., Schmidheini T.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN SEQUENCE OF 1-74 FROM N.A.
 RX MEDLINE=86230135; PubMed=2423850;
 RA Spevak W., Hartig A., Weindl P., Ruis H.;
 RT "Home control region of the catalase T gene of the yeast
 RT Saccharomyces cerevisiae.";
 RL Mol. Gen. Genet. 203:73-78 (1986).
 RN [5]
 RN SEQUENCE OF 405-409.
 RC STRAIN=ATCC 44827 / SKO2N;
 RX MEDLINE=97190279; PubMed=9038161;
 RA Norbeck J., Blomberg A.;
 RT "Metabolic and regulatory changes associated with growth of
 RT Saccharomyces cerevisiae in 1.4 M NaCl. Evidence for osmotic
 RT induction of glycerol dissimilation via the dihydroxyacetone
 RT pathway.";
 RL J. Biol. Chem. 272:5544-5554 (1997).
 CC -1- FUNCTION: Occurs in almost all aerobically respiring organisms and
 CC serves to protect cells from the toxic effects of hydrogen
 CC peroxide.
 CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
 CC -1- COFACTOR: Heme group.
 CC -1- SUBUNIT: Homotetramer.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- MISCELLANEOUS: This is one of two catalases in S.cerevisiae; the
 CC other is catalase A, which is the peroxisomal form.
 CC -1- SIMILARITY: Belongs to the catalase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, X04625; CA282828.1; -;
 DR EMBL, Z72873; CA97090.1; -;
 DR EMBL, M30256; AAA34540.1; -;
 DR PIR, S64383; CSBYT.
 DR HSSP, P04040; 1PAJ.
 DR Genmonline, 141400; -;
 DR SGD, S0003320; CTT1.
 DR GO, GO:0006950; P:response to stress; IDA.
 DR InterPro, IPR002226; Catalase.
 DR Pfam, PF000199; catalase; 1.
 DR PRINTS, PR00067; CATALASE.
 DR Prodom, PD000510; Catalase; 1.
 DR PROSITE, PS00437; CATALASE_1; 1.
 DR PROSITE, PS00438; CATALASE_2; 1.
 DR Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxidase;
 KM Multigene family.
 FT ACT_SITE 75 75 BY SIMILARITY.
 FT ACT_SITE 148 148 BY SIMILARITY.

FT METAL 362 362 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT CONFLICT 440 440 D -> V (IN REF. 1).
 FT CONFLICT 550 550 C -> G (IN REF. 1).
 SQ SEQUENCE 573 AA; 65741 MW; E03380543767377B CRC64;
 Query Match 43.3%; Score 42; DB 1; Length 573;
 Best Local Similarity 53.3%; Pred. No. 16;
 Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 3 LEAKICHOIEQFGD 17
 Db 191 LNPESIHQITFMFGD 205
 RESULT 13
 AGOI_ARATH STANDARD; PRT; 1048 AA.
 AC 004379; Q9LPE3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Argonaute protein.
 GN AGOI OR AT1G48410 OR F1A17.3 OR T1N15.2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OK NCBI_TaxID=3702;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia; TISSUE=leaf;
 RX MEDLINE=98090460; PubMed=9427751;
 RA Bohmert K., Camus I., Bellini C., Bouchet D., Caboche M., Benning C.;
 RT "AGO1 defines a novel locus of Arabidopsis controlling leaf
 RT development.";
 RL EMBL J. 17:170-180 (1998).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldhym T.V., Feng J.-D., Pong B., Fujii C.Y.,
 RA Gull J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huzar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Malti R., Marziani A.,
 RA Miltschker J., Miranda M., Nguyen M., Nierman W.C., Osborne B.T.,
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter U.C., Davis R.W.,
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:816-820 (2000).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=22954850; PubMed=14593172;
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gujral M., Hansen N.F.,
 RA Hayashizaki Y., Johnson-Hopson C., Huan V.W., Iida K., Karnes M.,
 RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamuya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,

```

RPSD MYCGE          STANDARD:          PRT:          497 AA.
ID   RPSD MYCGE
AC   P47491; Q49487;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DE   16-OCT-2001 (Rel. 40, Last annotation update)
DE   RNA polymerase sigma factor rpoD (Sigma-A).
GN   RPOD OR SIGA OR MG249.
OS   Mycoplasma genitalium.
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 33530 / G-37;
RA   Frazer C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA   Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA   Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Furmann J.L.,
RA   Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA   Tomb J.F., Dougherty B.A., Bort K.F., Hu P.-C., Luetter T.S.,
RA   Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT   "The minimal gene complement of Mycoplasma genitalium.";
RL   Science 270:397-403 (1995).
RN   [2]
RP   SEQUENCE OF 223-323 FROM N.A.
RC   STRAIN=ATCC 33530 / G-37;
RA   MEDLINE=92051396; PubMed=1945886;
RA   Peterson S.N., Schramm N., Hu P.-C., Bort K.F., Hutchison C.A. III,
RT   "A random sequencing approach for placing markers on the physical map
RT   of Mycoplasma genitalium.";
RL   Nucleic Acids Res. 19:6027-6031 (1991).
CC   -1- FUNCTION: The sigma factor is an initiation factor that promotes
CC   attachment of the RNA polymerase to specific initiation sites and
CC   then is released. This is the primary sigma-factor of this
CC   bacteria (by similarity).
CC   -1- SIMILARITY: Belongs to the sigma-70 factor family.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; U39703; AAC71469.1; -
DR   EMBL; X61535; CAB98133.1; -
DR   PIR; E64227; E64227.
DR   HSSP; P00579; 1SIG.
DR   TIGR; MG249; -.
DR   InterPro; IPR009043; RNA_pol_sigma.
DR   InterPro; IPR007627; Sigma70_r2.
DR   InterPro; IPR007624; Sigma70_r3.
DR   InterPro; IPR009943; Sigma70.
DR   Pfam; PF04542; sigma70_r2; 1.
DR   Pfam; PF04539; sigma70_r3; 1.
DR   PRINTS; PR00046; SIGMA70FCT.
DR   PROSITE; PS00715; SIGMA70_1; 1.
DR   PROSITE; PS00716; SIGMA70_2; 1.
DR   Transcription regulation; Sigma factor; DNA-directed RNA polymerase;
KW   DNA-binding; Complete proteome.
FT   DOMAIN 274 287
FT   DNA BIND 451 470
FT   CONFLICT 224 232
FT   CONFLICT 497 576
FT   SEQUENCE 497 AA; 57661 MW; ADE06CA68F59A5 CRC64;
SQ

```

```

Query Match          43.3%; Score 42; DB 1; Length 497;
Best Local Similarity 47.1%; Pred. No. 13;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
QY 2 ALEAKICHQIEYQFGDF 18
DB 382 SLDKTVGHDEBSQFGDF 398

```

```

RESULT 11
RPSD MYCPN          STANDARD:          PRT:          499 AA.
ID   RPSD MYCPN
AC   P78022;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DE   16-OCT-2001 (Rel. 40, Last annotation update)
DE   RNA polymerase sigma factor rpoD (Sigma-A).
GN   RPOD OR SIGA OR MPN352 OR MP484.
OS   Mycoplasma pneumoniae.
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2104;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29342 / M129;
RA   MEDLINE=97105885; PubMed=8948633;
RA   Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA   Herrmann R.;
RT   "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT   pneumoniae.";
RL   Nucleic Acids Res. 24:4420-4449 (1996).
CC   -1- FUNCTION: The sigma factor is an initiation factor that promotes
CC   attachment of the RNA polymerase to specific initiation sites and
CC   then is released. This is the primary sigma-factor of this
CC   bacteria (by similarity).
CC   -1- SIMILARITY: Belongs to the sigma-70 factor family.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; AE000047; AAB96132.1; -
DR   PIR; S73810; S73810.
DR   HSSP; P00579; 1SIG.
DR   InterPro; IPR009043; RNA_pol_sigma.
DR   InterPro; IPR007627; Sigma70_r2.
DR   InterPro; IPR007624; Sigma70_r3.
DR   InterPro; IPR009943; Sigma70.
DR   Pfam; PF04542; sigma70_r2; 1.
DR   Pfam; PF04539; sigma70_r3; 1.
DR   PRINTS; PR00046; SIGMA70FCT.
DR   PROSITE; PS00715; SIGMA70_1; 1.
DR   PROSITE; PS00716; SIGMA70_2; 1.
DR   Transcription regulation; Sigma factor; DNA-directed RNA polymerase;
KW   DNA-binding; Complete proteome.
FT   DOMAIN 276 289
FT   DNA BIND 453 472
FT   CONFLICT 499 576
FT   SEQUENCE 499 AA; 57796 MW; 7981C3BDE7C1E1 CRC64;
SQ

```

```

Query Match          43.3%; Score 42; DB 1; Length 499;
Best Local Similarity 47.1%; Pred. No. 14;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
QY 2 ALEAKICHQIEYQFGDF 18
DB 384 SLDKTVGHDEBSQFGDF 400

```

```

RESULT 12
CATT YEAST          STANDARD:          PRT:          573 AA.
ID   CATT YEAST
AC   P06115;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Catalase T (EC 1.11.1.6).
GN   CTT1 OR YGR088W.

```

RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mikhlin N.V., Modary C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
 RA Palazzolo M., Pictan G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheller F., Shen H.,
 RA Shue B.C., Sider-Kimam I., Simpson M., Skupski M.P., Smith T.,
 RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weissbrock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: May be involved in transcription termination by RNA
 polymerase III. Binds RNA and DNA. Binds to precursors of RNA
 polymerase III transcripts. May play a specialized role during fly
 development.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DEVELOPMENTAL STAGE: Expressed throughout embryonic, larval,
 pupal, and adult development. Expression throughout the embryo is
 followed by a restricted pattern of mesodermal expression that is
 later confined to the visceral mesoderm, gonads, gut, and salivary
 glands.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -1- SIMILARITY: TO VERTERATE PROTEIN LA.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, U07652; AAA20518.1; -;
 DR EMBL, L32988; AAA21776.1; -;
 DR EMBL, AE003666; AAF5385.1; -;
 DR PIR, A53773; A53773.
 DR PIR, A53781; A53781.
 DR FlyBase, FBgn0011638; La.
 DR GO, GO:0008098; F:5S rRNA primary transcript binding; IDA.
 DR GO, GO:0003723; F:RNA binding; NAS.
 DR InterPro, IPR002344; Lupus La.
 DR InterPro, IPR006530; Lupus La dom.
 DR InterPro, IPR005054; RNA_rec_mot.
 DR Pfam, PF05383; La; 1.
 DR Pfam, PF00076; xtm; 1.
 DR PRINTS, PR00302; LUPUSLA.
 DR SMART, SM00715; LA; 1.
 DR SMART, SM00360; RRM; 1.
 DR PROSITE, PS50102; RRM; 1.
 DR PROSITE, PS00303; RRM_RNP_1; 1.
 DR RNA-binding; Nuclear protein; DNA-binding.
 DR KEGG, K04914; RRM; 1.
 DR DOMAIN, 149 234
 FT CONFLICT 169 169 A -> T (IN REF. 1).
 FT CONFLICT 182 183 KH -> NS (IN REF. 1).
 FT CONFLICT 283 283 A -> R (IN REF. 1).
 FT CONFLICT 329 329 K -> N (IN REF. 1).
 SQ SEQUENCE 390 AA; 44884 MW; A8099288B90446A5 CRC64;

Query Match 44.3%; Score 43; DB 1; Length 390;
 Best Local Similarity 57.1%; Pred. No. 7;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGD 17
 DB 51 ERAIRQVEYYFGD 64

RESULT 9
 ID LBP_RABIT STANDARD; PRT; 482 AA.
 AC P17454;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Lipopolysaccharide-binding protein precursor (LBP).
 GN LBP.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90385281; PubMed=2402637;
 RA Schumann R.R., Leong S.R., Flagg G.W., Gray P.W., Wright S.D.,
 RA Mathison J.C., Tobias P.S., Ulevitch R.J.;
 RT "Structure and function of lipopolysaccharide binding protein.";
 RL Science 249:1429-1431(1990).
 RN [2]
 RP SEQUENCE OF 27-66.
 RC TISSUE=Serum;
 RX MEDLINE=86306528; PubMed=2427635;
 RA Tobias P.S., Soldau K., Ulevitch R.J.;
 RT "Isolation of a lipopolysaccharide-binding acute phase reactant from
 rabbit serum.";
 RL J. Exp. Med. 164:777-793(1986).
 CC -1- FUNCTION: Binds to the lipid moiety of bacterial
 lipopolysaccharides (LPS), a glycolipid present in the outer
 membrane of all Gram-negative bacteria. The LBP/LPS complex seems
 to interact with the CD14 receptor.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to the BPI/LBP/Plunc superfamily. BPI/LBP
 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, M35534; AAA9235.1; -;
 DR PIR, B35843; B35843.
 DR HSSP, P17213; LBPI.
 DR InterPro, IPR001124; LBP_BPI_CETP.
 DR Pfam, PF01273; LBP_BPI_CETP; 1.
 DR Pfam, PF02886; LBP_BPI_CETP_C; 1.
 DR SMART, SM00328; BPI; 1.
 DR SMART, SM00329; BPI2; 1.
 DR PROSITE, PS00400; LBP_BPI_CETP; 1.
 DR KEGG, K04914; LBP; 1.
 DR LipoLipid transport; Antibiotic; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 482
 FT CARBOHYD 301 301 LIPOLYSACCHARIDE-BINDING PROTEIN.
 FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 57 57 E -> G (IN REF. 2).
 FT CONFLICT 63 63 S -> F (IN REF. 2).
 SQ SEQUENCE 482 AA; 54001 MW; 628A5E0A647200C2 CRC64;

Query Match 43.3%; Score 42; DB 1; Length 482;
 Best Local Similarity 80.0%; Pred. No. 13;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LEAKICHOIE 12
 DB 194 LESKICHOIE 203

RESULT 10

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; X68817; CAA48715.1; -

DR PIR; S33818; S33818.

DR InterPro; IPR002344; Lupus_La.

DR InterPro; IPR006630; Lupus_La_dom.

DR InterPro; IPR00504; RNA_rec_mot.

DR Pfam; PF05383; La; 1.

DR Pfam; PF00076; rrm; 1.

DR PRINTS; PR00302; LUPUSLA.

DR SMART; SM00715; LA; 1.

DR SMART; SM00360; RRM; 1.

DR PROSITE; PS50102; RRM; 1.

DR PROSITE; PS00030; RRM_RNP_1; 1.

DR RNA-binding; Nuclear protein; Phosphorylation.

DR RNA-binding; Nuclear protein; RNA-BINDING (RRM).

FT DOMAIN 111 203 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

FT DOMAIN 196 212 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

FT DOMAIN 316 332 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

SQ SEQUENCE 428 AA; 48864 MW; AEB3A38B7D2E3EC3 CRC64;

Query Match 59.8%; Score 58; DB 1; Length 428;
Best Local Similarity 71.4%; Pred. No. 0.019;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 EAKICHQIEYQFGD 17
Db 14 DTKICQIEYQFGD 27

RESULT 7
LA AEDAL STANDARD; PRT; 383 AA.
ID LA AEDAL
AC 026457;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
OS Aedes albopictus (Forest day mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.
NCBI_TaxID=7160;

RN (1)
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96135233; PubMed=8551578;
RA Pardigon N., Straus J.H.;
RT "Mosquito homolog of the La autoantigen binds to Sindbis virus RNA.";
RL J. Virol. 70:1173-1181(1996).

CC -1- FUNCTION: May be involved in transcription termination by RNA
polymerase III. Binds RNA and DNA. Binds to the 3' end of the
minus strand of Sindbis virus RNA. This may be significant for
Sindbis virus RNA replication.
CC -1- SUBCELLULAR LOCATION: Nuclear. Primarily nuclear, but significant
amounts are present in the cytoplasm.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -1- SIMILARITY: TO VERTERATE PROTEIN LA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC -----

DR EMBL; S80954; AAB35931.1; -

DR InterPro; IPR002344; Lupus_La.

DR InterPro; IPR006630; Lupus_La_dom.

DR InterPro; IPR00504; RNA_rec_mot.

DR Pfam; PF05383; La; 1.

DR Pfam; PF00076; rrm; 1.

DR PRINTS; PR00302; LUPUSLA.

DR SMART; SM00715; LA; 1.

DR SMART; SM00360; RRM; 1.

DR PROSITE; PS50102; RRM; 1.

DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.

KW RNA-binding; Nuclear protein; DNA-binding.

FT DOMAIN 141 228 RNA-BINDING (RRM).

FT DOMAIN 196 212 RNA-BINDING (RRM).

SQ SEQUENCE 383 AA; 44430 MW; 4E5C8F21C40F452 CRC64;

Query Match 48.5%; Score 47; DB 1; Length 383;
Best Local Similarity 60.0%; Pred. No. 1.4;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 LEAKICHQIEYQFGD 17
Db 43 LEASTIRQIEYQFGD 57

RESULT 8
LA DROME STANDARD; PRT; 390 AA.
ID LA DROME
AC P40796; Q24375; Q9VIN2;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
GN LA OR CG10922.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;

RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Cancon-S; TISSUE=Ovary;
RX MEDLINE=94309632; PubMed=8035794;
RA Bai C., Li Z., Tollas P.P.;
RT "Developmental characterization of a Drosophila RNA-binding protein
homologous to the human systemic lupus erythematosus-associated
La/SS-B autoantigen";
RL Mol. Cell. Biol. 14:5123-5129(1994).

RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=94309661; PubMed=8035818;
RA Yoo C.J., Molin S.L.;
RT "La proteins from Drosophila melanogaster and Saccharomyces
cerevisiae: a yeast homolog of the La autoantigen is dispensable for
growth";
RL Mol. Cell. Biol. 14:5412-5424(1994).

RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Mortman J.R., Yandell M.D., Zhang O., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,
Man R.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Bailly R.M., Basu A., Bakendell J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolashov S.,
Borkova D., Botchan M.R., Bouck J., Brooksstein P., Brotler P.,
Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
Rusler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D.A., Helman T.J., Hernandez J.R., Houck J.,
Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; L00993; AAA39415.1; -
 DR EMBL; BC003820; AA03820.1; -
 DR EMBL; Y07951; CAA69249.1; -
 DR MGI; MGI:98423; Ssb.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR InterPro; IPR002344; Lopus_La.
 DR InterPro; IPR006630; Lopus_La.
 DR InterPro; IPR00504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rtm; 1.
 DR SMART; SMART:SM00360; RRM; 1.
 DR SMART; SMART:SM00715; LA; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS0030; RRM_RNP_1; 1.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 KW DOMAIN 111 187 RNA-BINDING (RRM).
 FT SEQUENCE 415 AA; 47756 MW; 2D75197692FDC933 CRC64;

Query Match 89.7%; Score 87; DB 1; Length 415;
 Best Local Similarity 94.1%; Pred. No. 1.8e-07;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALEAKICHOIEYFGDF 18
 Db 12 ALEAKICHOIEYFGDF 28

RESULT 5
 ID LAB_XENLA STANDARD; PRT; 427 AA.
 AC P28049;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lopus La protein homolog B (La ribonucleoprotein B) (La autoantigen
 DE homolog B).
 GN LAB1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocyte;
 RX MEDLINE=93287095; PubMed=8510143;
 RA Scherly D., Stutz F., Lin-Marg N., Clarkson S.G.;
 RT "La proteins from Xenopus laevis. cDNA cloning and developmental
 RT expression.";
 RL J. Mol. Biol. 231:196-204(1993).
 CC -!- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' terminus of virtually all nascent
 CC polymerase III transcripts (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- DEVELOPMENTAL STAGE: Barely detectable in stage I/II oocytes,
 CC accumulate in stage III/IV oocytes, then exhibit a roughly
 CC constant steady state level in mature oocytes, eggs, and early
 CC embryos.
 CC -!- PTM: Phosphorylated (Probable).
 CC -!- MISCELLANEOUS: There are two forms of La, LaA and LaB, in xenopus.
 CC -!- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
 CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X68818; CAA48716.1; -
 DR PIR; S33817; S33817.
 DR InterPro; IPR002344; Lopus_La.
 DR InterPro; IPR006630; Lopus_La.
 DR InterPro; IPR00504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rtm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SMART:SM00715; LA; 1.
 DR SMART; SMART:SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS0030; RRM_RNP_1; FALSE NEG.
 KW RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 110 202 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 315 331
 FT SEQUENCE 427 AA; 48995 MW; 45F3146F8934A355 CRC64;

Query Match 63.9%; Score 62; DB 1; Length 427;
 Best Local Similarity 73.3%; Pred. No. 0.0039;
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LEAKICHOIEYFGD 17
 Db 12 LDTKICEQIEYFGD 26

RESULT 6
 ID LAB_XENLA STANDARD; PRT; 428 AA.
 AC P28048;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lopus La protein homolog A (La ribonucleoprotein A) (La autoantigen
 DE homolog A).
 GN LAB1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocyte;
 RX MEDLINE=93287095; PubMed=8510143;
 RA Scherly D., Stutz F., Lin-Marg N., Clarkson S.G.;
 RT "La proteins from Xenopus laevis. cDNA cloning and developmental
 RT expression.";
 RL J. Mol. Biol. 231:196-204(1993).
 CC -!- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' terminus of virtually all nascent
 CC polymerase III transcripts (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- DEVELOPMENTAL STAGE: Barely detectable in stage I/II oocytes,
 CC accumulate in stage III/IV oocytes, then exhibit a roughly
 CC constant steady state level in mature oocytes, eggs, and early
 CC embryos.
 CC -!- PTM: Phosphorylated (Probable).
 CC -!- MISCELLANEOUS: There are two forms of La, LaA and LaB, in xenopus.
 CC -!- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
 CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

DR PROSITE; PS00030; RRM_RNP_1; 1.
 KM Systemic lupus erythematosus; RNA-binding; Phosphorylation;
 NM Nuclear protein.
 FT DOMAIN 111 187 RNA-BINDING (RRM).
 FT MOD_RES 366 366 PHOSPHORYLATION (BY CK2).
 SQ SEQUENCE 408 AA; 46837 MW; EC153C15F9187PC4 CRC64;

Query Match 93.8%; Score 91; DB 1; Length 408;
 Best Local Similarity 94.4%; Pred. No. 3.5e-08;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYFGDF 18
 |||||
 DB 11 AALEAKICHQIEYFGDF 28

RESULT 3
 ID LA RAT STANDARD; PRT; 415 AA.
 AC P36566;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
 DE homolog).
 GN SSB OR SS-B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93246255; PubMed=7916708;
 RA Semsei I., Troester H., Bartsch H., Schwemmler M., Igloi G.L.,
 RA Bachmann M.;
 RT "Isolation of rat cDNA clones coding for the autoantigen SS-B/La:
 RT detection of species-specific variations.";
 RL Gene 126:265-268(1993).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' terminus of virtually all nascent
 CC polymerase III transcripts. It is associated with precursor forms
 CC of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S,
 CC and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: Phosphorylated (By similarity).
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X67859; CAA48043.1; -.
 CC PIR; JCI494; JCI494.
 CC DR InterPro; IPR002344; Lupus_La.
 CC DR InterPro; IPR006630; Lupus_La_dom.
 CC DR InterPro; IPR000504; RNA_rec_mot.
 CC Pfam; PF00383; La; 1.
 CC DR Pfam; PF00076; rrm; 1.
 CC DR PRINTS; PR00302; LUPUSLA.
 CC SMART; SM00715; La; 1.
 CC SMART; SM00360; RRM; 1.
 CC DR PROSITE; PS50102; RRM; 1.
 CC DR PROSITE; PS00030; RRM_RNP_1; 1.
 CC KM RNA-binding; Nuclear protein; Phosphorylation.
 CC FT DOMAIN 111 187 RNA-BINDING (RRM).
 CC SQ SEQUENCE 415 AA; 47777 MW; 033FD9CC18475F98 CRC64;

Query Match 93.8%; Score 91; DB 1; Length 415;
 Best Local Similarity 94.4%; Pred. No. 3.5e-08;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYFGDF 18
 |||||
 DB 11 AALEAKICHQIEYFGDF 28

RESULT 4
 ID LA MOUSE STANDARD; PRT; 415 AA.
 AC P32067;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
 DE homolog).
 GN SSB OR SS-B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93203630; PubMed=8454877;
 RA Topfer F., Gordon T., McCluskey J.;
 RT "Characterization of the mouse autoantigen La (SS-B). Identification
 RT of conserved RNA-binding motifs, a putative ATP binding site and
 RT reactivity of recombinant protein with poly(U) and human
 RT autoantibodies.";
 RL J. Immunol. 150:3091-3100(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=FVB/N; TISSUE=Mammary gland;
 RC MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Bhat N.K.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ushed T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mallya S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schercher A., Schein J.E., Jones S.J.M., Maira M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE OF 1-11 FROM N.A.
 RA Groelz D., Bachmann M.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' terminus of virtually all nascent
 CC polymerase III transcripts. It is associated with precursor forms
 CC of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S,
 CC and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: Phosphorylated (By similarity).
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

FT DOMAIN 111 187 RNA-BINDING (RRM).
 SQ SEQUENCE. 404 AA; 46534 MW; 4EE30B5C262AD6A1 CRC64;
 Query Match 93.8%; Score 91; DB 1; Length 404;
 Best Local Similarity 94.4%; Pred. No. 3,4e+08;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHQIRYQGF 18
 |||||
 Db 11 AALEAKICHQIRYQGF 28

RESULT 2
 ID LA_HUMAN STANDARD; PRT; 408 AA.
 AC P05455;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein (Sjogren syndrome type B antigen) (SS-B) (La ribonucleoprotein) (La autoantigen).
 GN SSB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_taxid:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89202037; PubMed=2468131;
 RA Chan E.K.L., Sullivan K.F., Tan E.M.;
 RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences for RNA-binding.";
 RL Nucleic Acids Res. 17:2233-2244 (1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89053970; PubMed=3192525;
 RA Chambers J.C., Kenan D., Martin B.J., Keene J.D.;
 RT "Genomic structure and amino acid sequence domains of the human La autoantigen.";
 RL J. Biol. Chem. 263:18043-18051 (1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta; and Skeletal muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Abramson R.D., Mullen L.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stachleir M., Soares W.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Frange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fanev J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield J.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
 RA Schercher A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [4]
 RP SEQUENCE OF 54-408 FROM N.A.
 RX MEDLINE=88199081; PubMed=2452201;
 RA Sturgees A.D., Peterson M.G., McNeilage L.J., Whittingham S.,
 RA Coppel R.S.;
 RT "Characteristics and epitope mapping of a cloned human autoantigen La.";
 RL J. Immunol. 140:3212-3218 (1988).
 RN [5]

RP SEQUENCE OF 54-97 FROM N.A.
 RX MEDLINE=8516283; PubMed=3856888;
 RA Chambers J.C., Keene J.D.;
 RT "Isolation and analysis of cDNA clones expressing human lupus La antigen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2115-2119 (1985).
 RN [6]
 RP FUNCTION.
 RX MEDLINE=89251617; PubMed=2470590;
 RA Gottlieb E., Steltz J.A.;
 RT "Function of the mammalian La protein: evidence for its action in transcription termination by RNA polymerase III.";
 RL EMBL J. 8:851-861 (1989).
 RN [7]
 RP PHOSPHORYLATION.
 RX MEDLINE=97207017; PubMed=9054510;
 RA Fan H., Sakulich A.L., Goodier J.L., Zhang X., Qin J., Marate R.J.;
 RT "Phosphorylation of the human La antigen on serine 366 can regulate recycling of RNA polymerase III transcription complexes.";
 RL Cell 88:707-715 (1997).
 RN [8]
 RP INTERACTION WITH DDX15.
 RX MEDLINE=22346609; PubMed=12458796;
 RA Fouraux M.A., Kolkman M.J.M., Van der Heijden A., De Jong A.S.,
 RA Van Venrooij W.J., Pruijn G.J.M.;
 RT "The human La (SS-B) autoantigen interacts with DDX15/hPrp43, a putative DEAD-box RNA helicase.";
 RL RNA 8:1428-1443 (2002).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' termini of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S, and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE C-TERMINAL PART OF THE PROTEIN.
 CC -1- PTM: The N-terminus is blocked.
 CC -1- DISEASE: Sera from patients with systemic lupus erythematosus often contain antibodies that react with the normal cellular La protein as if this antigen was foreign.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X13697; CAA31985.1; -;
 CC EMBL: J04205; AAA51885.1; -;
 CC EMBL: BC001289; AAH01289.1; -;
 CC EMBL: BC020818; AAH20818.1; -;
 CC PIR: A31888; A31888.
 CC GeneW: HGNC:11316; SSB.
 CC MIM: 109090; -;
 CC GO: GO:0030529; C:ribonucleoprotein complex; TAS.
 CC GO: GO:0003729; F:mRNA binding; TAS.
 CC GO: GO:0000049; F:tRNA binding; TAS.
 CC GO: GO:0008334; P:histone mRNA metabolism; TAS.
 CC GO: GO:0006400; P:tRNA modification; TAS.
 CC InterPro: IPR002344; Lupus La.
 CC InterPro: IPR006630; Lupus La dom.
 CC InterPro: IPR000504; RNA_rec_mot.
 CC Pfam: PF05383; La; 1.
 CC Pfam: PF00076; trm; 1.
 CC PRINTS: PR00302; LUPUSLA.
 CC SMART: SM00715; La; 1.
 CC SMART: SM00360; RRM; 1.
 CC PROSITE: PS50102; RRM; 1.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:01:41; Search time 7.24022 Seconds
(without alignments)
129.452 Million cell updates/sec

Title: US-09-836-073-12

Perfect score: 97
Sequence: 1 AALEKICHQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	93.8	404	1	LA_BOVIN
2	91	93.8	408	1	LA_HUMAN
3	91	93.8	415	1	LA_RAT
4	87	89.7	415	1	LA_MOUSE
5	62	63.9	427	1	LA_XENLA
6	58	59.8	428	1	LA_XENLA
7	47	48.5	383	1	LA_AEDAL
8	43	44.3	390	1	LA_DROME
9	42	43.3	482	1	LA_RABIT
10	42	43.3	487	1	RPSD_MYCGE
11	42	43.3	499	1	RPSD_MYCPN
12	42	43.3	573	1	CATT_YEAST
13	42	43.3	1048	1	AGOI_ARATH
14	42	42.3	506	1	TCMO_PPCRC
15	41	42.3	902	1	ATMA_SALTY
16	41	42.3	988	1	PINW_ARATH
17	40	41.2	357	1	CARA_BUCBP
18	40	41.2	490	1	IFT4_HUMAN
19	40	41.2	591	1	Y875_CHLTR
20	40	41.2	1107	1	MY1B_MOUSE
21	40	41.2	1136	1	MY1B_RAT
22	40	41.2	2556	1	NTC1_HUMAN
23	39	40.2	158	1	YV02_HUMAN
24	39	40.2	231	1	MTNA_BACSU
25	39	40.2	310	1	PP12_TOBAC
26	39	40.2	461	1	FUMC_RICPR
27	39	40.2	501	1	XYLB_LACPE
28	39	40.2	552	1	VP33_ARATH
29	39	40.2	812	1	PLMN_MOUSE
30	39	40.2	997	1	AGOL_ARATH
31	39	40.2	1174	1	KPOL_COCH
32	39	40.2	1433	1	VGLM_BUNYW
33	39	40.2	2483	1	COA2_HUMAN

34	38.5	39.7	1036	1	OGT1_HUMAN
35	38.5	39.7	1036	1	OGT1_RAT
36	38	39.2	298	1	LAH1_SCHPO
37	38	39.2	479	1	PR12_ARATH
38	38	39.2	487	1	GARB_CHICV
39	38	39.2	505	1	TCMO_ARATH
40	38	39.2	505	1	TCMO_CICAR
41	38	39.2	505	1	TCMO_GLYEC
42	38	39.2	505	1	TCMO_HELTV
43	38	39.2	505	1	TCMO_PEA
44	38	39.2	505	1	TCMO_PHAU
45	38	39.2	505	1	TCMO_ZINEL

ALIGNMENTS

RESULT 1
LA_BOVIN STANDARD; PRT; 404 AA.
ID LA_BOVIN
AC P10981;
DT 01-UTL-1989 (Rel. 11, Created)
DT 01-UTL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
GN SSB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Plutitary; PubMed=2468131;
RX MEDLINE=89202037; Pubmed=2468131;
RA Chan E.K.L., Sullivan K.F., Tan E.M.;
RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences for RNA-binding".
RT Nucleic Acids Res. 17:2233-2244(1989).
RL
CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' termin of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S, and 7-2 RNAs.
CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE C-TERMINAL PART OF THE PROTEIN.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch) or send an email to license@isb-sib.ch).
CC EMBL; X13698; CAA31986.1; -
DR 2IR; S03849; S03849.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
KM RNA-binding; Nuclear protein; Phosphorylation.

This Page Blank (uspto)

Best Local Similarity 47.1%; Pred. No. 39;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 ALEAKICHQIEYQFGDF 18

Db 382 SLDKTVGHDEESQFGDF 398

RESULT 15

transcription initiation factor sigma A - Mycoplasma pneumoniae (strain ATCC 29342)

N;Alternate names: hypothetical protein H91_orf499

C;Species: Mycoplasma pneumoniae

A; Variety: ATCC 29342

```
C:\Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1997
```

C;Accession: S73810

R.; Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.; Nucleic Acids Res 34 4420-4449 1996

Nucleic Acids Res. 24, 4420-4449, 1996

A:Title: Complete sequence analysis of the genome of the bacterium *Mycobacterium*

A: Accession: S73810
A: Reference number: S/332/; MUID:9/105885; PMID:8946833

A;Accession: S/3810

A:Molecule type: DNA
A>Status: preliminary; nucleic acid sequence not known; translation not known

A:Residues: 1-499 <H

A; Cross-references: EME

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996

C:\Genetics:

A;Gene: *sigA*

A; Genetic co

C;Superfamily: trans

C;Keywords: DNA binding; sigma factor; transcription initiation

F;254-487/Domain: transcription factor sigma katF hcr

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Query Match	Score	DB	Length
43.3%	42	2	499

Best Local Similarity 47.1%; Pred. No. 39;

Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0

QY 2 ALEAKICHQIEYQFGDF 18

Db 384 SLDKTVGHDEESQFGDF 400

Search completed: September 10, 2004, 18:02:30
Job time : 16.0782 secs

Job time : 16.0782 secs

Query March	44.3%	Score 43;	DB 2;	Length 390;
Best Local Similarity	57.1%	Pred. No. 21;		
Matches	8;	Conservative	1;	Mismatches 5;
				Indels 0;
				Gaps 0;
OY	4	EAKICHQLEYPFGD	17	
Db	51	ERAIRQVEYYFGD	64	

Query Match	44.3%	Score 43	DB 2	Length 1131
Best Local Similarity	70.0%	Pred. No. 60		
Matches 7, Conservative		1; Mismatches 2	Indels 0	Gaps 0

A;Gene: CESP:C30H6.2

Query Match	43.3%	Score 42	DB 2	Length 360
Best Local	46.2%	Pred. NC	28	
Matches	6	Conservative	4	Mismatches 3; Indels 0; Gaps 0;

c;Keywords: acute phase
Query Match 43 38: Score 43: DB 2

QY	3 LEAKCHOIE 12	43.3%	Score 42;	DB 2;	Length 483;
	Query Match	43.3%	Score 42;	DB 2;	Length 483;
	Best Local Similarity	80.0%	Pred. No. 38;		
	Matches 8;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0.

E64227

N:Alternate names: RNA polymerase sigma-A factor
C:Species: Mycoplasma genitalium

C;Accession: E64227

M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Sauder, D.M.; Phillips, C.A.; Merrick, J.

Science 270, 397-403

A:Reference number: A64200; MUID:96026346; PMID:7569993

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Residues: 1-497 <TIGR;

A; Experimental source: strain G-37

A;Genetic code: SGC3

Keywords: DNA binding; sigma factor; transcription initiation

100

Query Match

43.3%; Score 42; DB 2; Length 497;

A:Reference number: S03848; MUID:89202037; PMID:2468131
 A:Accession: S03848
 A:Molecule type: mRNA
 A:Residues: 1-408 <CH2>
 A:Cross-references: EMBL:X13697; NID:g36414; PIDN:CAA31985.1; PID:g36415
 R:Chambers, J.C.; Keene, J.D.
 Proc. Natl. Acad. Sci. U.S.A. 82, 2115-2119, 1985
 A:Title: Isolation and analysis of cDNA clones expressing human lupus La antigen.
 A:Reference number: A22956; MUID:85166283; PMID:3856888
 A:Accession: A22956
 A:Molecule type: mRNA
 A:Residues: 45-97, 'LK', <CH3>
 A:Cross-references: GB:J04205
 A:Note: this sequence has been revised in reference A31888
 R:Nyman, U.; Ringertz, N.R.; Petersen, I.
 Immunol. Lett. 22, 65-72, 1989
 A:Title: Demonstration of an amino terminal La epitope recognized by human anti-La sera.
 A:Reference number: A61051; MUID:89379261; PMID:2476379
 A:Accession: A61051
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-19, 'E', '21-47' <NYM>
 R:Sturgess, A.D.; Peterson, M.G.; McNeillage, L.J.; Whittingham, S.; Coppel, R.L.
 J. Immunol. 140, 3212-3218, 1988
 A:Title: Characteristics and epitope mapping of a cloned human autoantigen La.
 A:Reference number: S11013; MUID:8819081; PMID:2452201
 A:Accession: S11013
 A:Molecule type: mRNA
 A:Residues: 'E', 55-287, 'V', 289-408 <STU>
 A:Cross-references: EMBL:M20328; NID:g337456; PIDN:AAA36577.1; PID:g337457
 R:Kobasaka, H.; Yamamoto, K.; Fujii, H.; Miura, H.; Miyasaka, N.; Nishioaka, K.; Miyamoto, J.
 Clin. Invest. 85, 1566-1574, 1990
 A:Title: Fine epitope mapping the human SS-B/La protein: Identification of a distinct an
 A:Reference number: I55553; MUID:90237237; PMID:1692037
 A:Accession: I55553
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 81-107 <RES>
 A:Cross-references: GB:M5261; NID:g338491; PIDN:AAA36652.1; PID:g338495
 A:Accession: I70205
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 174-224 <RE2>
 A:Cross-references: GB:M5263; NID:g338492; PIDN:AAA36653.1; PID:g338496
 A:Accession: I70206
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 279-342 <RE3>
 A:Cross-references: GB:M5262; NID:g338493; PIDN:AAA36654.1; PID:g338497
 C:Comment: This protein associates with a variety of small RNA molecules, most of which
 act as a transcription termination factor.
 C:Genetics:
 A:Gene: GDB:SSB
 A:Cross-references: GDB:125359; OMIM:109090
 A:Map position: 2
 A:Introns: 22/3; 57/2; 115/3; 151/3; 185/2; 209/2; 223/3; 264/3; 380/2
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:112-178/Domain: ribonucleoprotein repeat homology <RMM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:228-408/Domain: phosphorylated #status experimental <PHY>

Query Match 93.8%; Score 91; DB 1; Length 408;
 Best Local Similarity 94.4%; Pred. No. 2e-07;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 AALEAKICHQIEYFGDF 18
 |||||
 Db 11 AALEAKICHQIEYFGDF 28

RESULT 3

JC1494
 ribonucleoprotein La - rat
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 24-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 A:Accession: JC1494; S25145
 R:Semse, I.; Troester, H.; Bartsch, H.; Schwemle, M.; Igloi, G.L.; Bachmann, M.
 Gene 126, 265-268, 1993
 A:Title: Isolation of rat cDNA clones coding for the autoantigen SS-B/La: Detection of B
 A:Reference number: JC1494; MUID:93246255; PMID:7916708
 A:Accession: JC1494
 A:Molecule type: mRNA
 A:Residues: 1-415 <SEM>
 A:Cross-references: GB:X67859; NID:g55778; PIDN:CAA48043.1; PID:g55779
 A:Experimental source: liver
 C:Comment: This protein associates with a variety of small RNA molecules, most of which
 act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:112-178/Domain: ribonucleoprotein repeat homology <RMM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:227-415/Domain: phosphorylated #status predicted <PHY>

Query Match 93.8%; Score 91; DB 1; Length 415;
 Best Local Similarity 94.4%; Pred. No. 2e-07;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 AALEAKICHQIEYFGDF 18
 |||||
 Db 11 AALEAKICHQIEYFGDF 28

RESULT 4
 S33817
 ribonucleoprotein La.B - African clawed frog
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 A:Accession: S33817; S28544
 R:Schery, D.; Stutz, F.; Lin-Marg, N.; Clarkson, S.G.
 J. Mol. Biol. 231, 196-204, 1993
 A:Title: La proteins from Xenopus laevis. cDNA cloning and developmental expression.
 A:Reference number: S33817; MUID:93287095; PMID:8510143
 A:Accession: S33817
 A:Molecule type: mRNA
 A:Residues: 1-427 <SCH>
 A:Cross-references: EMBL:X68818; NID:g64875; PIDN:CAA48716.1; PID:g64876
 C:Comment: This protein associates with a variety of small RNA molecules, most of which
 act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:111-177/Domain: ribonucleoprotein repeat homology <RMM>
 F:112-117/Region: RNA-binding RNP2 motif
 F:150-157/Region: RNA-binding RNP1 motif
 F:227-427/Domain: phosphorylated #status predicted <PHY>

Query Match 63.9%; Score 62; DB 1; Length 427;
 Best Local Similarity 73.3%; Pred. No. 0.015;
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 3 LEAKICHQIEYFGDF 17
 |||||
 Db 12 LDRKICQIEYFGDF 26

RESULT 5
 S33818
 ribonucleoprotein La.A - African clawed frog
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 A:Accession: S33818; S28545

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:33:11; Search time 14.0782 Seconds
(without alignments)
122.988 Million cell updates/sec

Title: US-09-836-073-12

Perfect score: 97

Sequence: 1 AALEAKICHQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*

1: D1L1:*
2: P1R2:*
3: P1R3:*
4: P1R4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	93.8	404	1 S03849	ribonucleoprotein
2	91	93.8	408	1 A31888	ribonucleoprotein
3	91	93.8	415	1 JCI194	ribonucleoprotein
4	62	63.9	427	1 S33817	ribonucleoprotein
5	58	59.8	428	1 S33818	ribonucleoprotein
6	49	50.5	166	2 T32701	hypothetical prote
7	44	45.4	568	2 C82379	response regulator
8	44	45.4	752	2 G02273	Liv-1 protein - hu
9	43	44.3	390	2 A53773	La/SS-B homolog D-
10	43	44.3	390	2 A53781	ribonucleoprotein
11	43	44.3	1131	2 T15617	hypothetical prote
12	42	43.3	360	2 T15685	hypothetical prote
13	42	43.3	482	2 B35843	lipopolysaccharide
14	42	43.3	497	2 E64227	transcription init
15	42	43.3	499	2 S73810	transcription init
16	42	43.3	541	2 H71887	hypothetical prote
17	42	43.3	542	2 G64627	hypothetical prote
18	42	43.3	573	1 CSBT	catalase (BC 1.11.
19	42	43.3	1250	2 AF1482	hypothetical prote
20	41	42.3	120	2 AF1482	hypothetical prote
21	41	42.3	150	2 A84488	hypothetical prote
22	41	42.3	391	2 H89777	capsular polysacch
23	41	42.3	506	2 T14907	trans-cinnamate 4-
24	41	42.3	540	2 E75633	conserved hypothet
25	41	42.3	902	2 B57147	Mg2+-transporting
26	41	42.3	902	2 AH1058	Mg(2+) transport A
27	41	42.3	988	2 T52134	Zwille protein (im
28	40.5	41.8	386	2 H86870	hypothetical prote
29	40	41.2	157	2 S58067	probable olfactory

30	40	41.2	159	2 C96967	lactylglutacton 1
31	40	41.2	233	2 A81945	probable adenosyl
32	40	41.2	233	2 D81159	5-methylthioadenos
33	40	41.2	529	2 T00677	hypothetical prote
34	40	41.2	591	2 C71460	hypothetical prote
35	40	41.2	658	2 T19487	hypothetical prote
36	40	41.2	674	2 UC5104	transcription init
37	40	41.2	862	2 T01798	hypothetical prote
38	40	41.2	1078	2 B45439	myosin-I, Myr 1c (
39	40	41.2	1079	2 A45438	myosin I alpha cha
40	40	41.2	1094	2 S32404	myosin heavy chain
41	40	41.2	1107	2 C45439	myosin-I, Myr 1b (
42	40	41.2	1136	2 A45439	myosin I heavy cha
43	40	41.2	1217	2 H89863	hypothetical prote
44	40	41.2	2555	2 A40043	notch protein homo
45	39	40.2	147	2 E95015	glycosyl transfera

ALIGNMENTS

RESULT 1
S03849
ribonucleoprotein La - bovine
N/Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
C/Species: Bos primigenius taurus (cattle)
C/Date: 31-Dec-1990 #sequence_revision 26-May-1994 #ext_change 22-Jun-1999
C/Accession: S03849
R/Chan, E.K.V.; Sullivan, K.F.; Tan, E.M.
Nucleic Acids Res. 17, 2233-2244, 1989
A/Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences
A/Reference number: S03849; MUID:89202037; PMID:2468131
A/Accession: S03849
A/Molecule type: mRNA
A/Residues: 1-404 <CHA>
A/Cross-references: EMBL:X13698; NID:g755; PIDN:CAA31986.1; PID:g756
A/Note: Part of this sequence was confirmed by protein sequencing
C/Comment: This protein associates with a variety of small RNA molecules, most of which
act as a transcription termination factor.
C/Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
C/Keywords: blocked amino end; phosphoprotein; RNA binding
F/112-178/Domain: ribonucleoprotein repeat homology <RRM>
F/113-118/Region: RNA-binding RNP2 motif
F/151-158/Region: RNA-binding RNP1 motif
F/228-404/Domain: phosphorylated #status predicted <PHY>
Query Match 93.8%; Score 91; DB 1; Length 404;
Best Local Similarity 94.4%; Pred. No. 2e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AALEAKICHQIEYFGDF 18
DB 11 AALEAKICHQIEYFGDF 28
RESULT 2
A31888
ribonucleoprotein La - human
N/Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B; Sjogren syndrome antigen
C/Species: Homo sapiens (man)
C/Date: 21-May-1990 #sequence_revision 26-May-1994 #ext_change 22-Jun-1999
C/Accession: A31888; S03848; A22956; A61051; S11013; I55553; I70205; I70206; A31273
R/Chan, E.K.V.; Kenan, D.; Martin, B.O.; Keene, J.D.
J. Biol. Chem. 263, 18043-18051, 1988
A/Title: Genomic structure and amino acid sequence domains of the human La autoantigen.
A/Reference number: A31888; MUID:89053970; PMID:3192525
A/Accession: A31888
A/Molecule type: mRNA
A/Residues: 1-408 <CHA>
A/Cross-references: GB:J04205; NID:g178686; PIDN:AAAS1885.1; PID:g178687
R/Chan, E.K.V.; Sullivan, K.F.; Tan, E.M.
Nucleic Acids Res. 17, 2233-2244, 1989
A/Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences

This Page Blank (uspto)

Mon Sep 13 09:36:24 2004

us-09-836-073-4.rsp

Page 7

Search completed: September 10, 2004, 18:00:03
Job time : 42.229 secs

Best Local Similarity 66.7%; Pred. No. 18;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIETFFGD 17
DB 136 KIVNVEYFFSD 207

RESULT 13

ID Q8A8Y5 PRELIMINARY; PRT; 756 AA.
AC Q8A8Y5;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Conserved hypothetical protein.
GN BT1032.
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_Taxid=818;
RN (1)
RP .SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=1263928;
RA Xu J., Bjursell M.K., Hilmrod J., Deng S., Carmichael L.K.,
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis."
RL Science 299:2074-2076 (2003).
DR EMBL; AB016930; AA076139.1; -
DR InterPro; IPR005887; alpha man.
DR TIGRFAMs; TIGR01180; aman2_put; 1.
KW Hypothetical protein. Complete proteome.
SQ SEQUENCE 756 AA; 86621 MM; 3DBY906689B45D4 CRC64;

Query Match 47.1%; Score 48; DB 16; Length 756;
Best Local Similarity 38.9%; Pred. No. 26;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 QOOEAKICHOIETFFGD 18
DB 492 EOYDMSVSHALEYIADF 509

RESULT 14

ID Q81455 PRELIMINARY; PRT; 2349 AA.
AC Q81455;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
GN PFE0230W.
OS Plasmodium falciparum (Isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_Taxid=36329;
RN (1)
RP .SEQUENCE FROM N.A.
RC Devlin K., Baker S., Davies P., Mungal K., Berriman M., Pain A.,
RA Hall N., Bowman S., Churcher C., Quail M., Barrett B.,
RT Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
[2]
RP .SEQUENCE FROM N.A.
RC MEDLINE=2255708; PubMed=12368867;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark C., Corton C.,
RA Croft A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Felwell T., Goble A., Goodhead I., Gilliam K., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holtroyd S., Horrocks P.,
RA Humphrey S., Jagels K., James K.D., Johnson D., Kethornu A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,

RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulten J.E., Craig A., Newbold C., Barrett B.G.
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13."
RL Nature 419:527-531 (2002).
DR EMBL; AL929351; CAD51412.1; -
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR InterPro; IPR000345; C:cyt heme BS.
DR InterPro; IPR002114; Hpr_Serp_S.
DR InterPro; IPR007087; Znf_C2H2.
DR PROSITE; PS00190; CYTOCHROME_C; 4.
DR PROSITE; PS00589; PTS_HPR_SER; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
KW Hypothetical protein.
SQ SEQUENCE 2349 AA; 291035 MM; FC33BFCCED94D563 CRC64;

Query Match 47.1%; Score 48; DB 5; Length 2349;
Best Local Similarity 38.9%; Pred. No. 79;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 QOOEAKICHOIETFFGD 18
DB 853 OEHVREMCPCPEYFTNY 870

RESULT 15

ID Q87T57 PRELIMINARY; PRT; 329 AA.
AC Q87T57;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Putative lipid A biosynthesis (kdo2-(lauroyl)-lipid IVA
acylttransferase.
GN VP0213.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_Taxid=670;
RN (1)
RP .SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Nishijima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.,
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
distinct from that of V. cholerae."
RL Lancet 361:743-749 (2003).
DR EMBL; AP005073; BAC58476.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0008415; F:acylttransferase activity; IEA.
DR GO; GO:0009244; F:lipopolysaccharide core region biosynthesis; IEA.
DR InterPro; IPR004960; LipA_acylttrans.
DR Pfam; PF03279; Lip_A_acylttrans; 1.
KW Complete proteome.
SQ SEQUENCE 329 AA; 37996 MM; 3F3C6C41D91B5D0C CRC64;

Query Match 46.6%; Score 47.5; DB 16; Length 329;
Best Local Similarity 44.4%; Pred. No. 13;
Matches 8; Conservative 7; Mismatches 2; Indels 1; Gaps 1;

QY 1 QOOEAKICHOIETFFGD 17
DB 275 EOEDARTCNLEAYFFVD 292

AC 044678;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
 DE Hypothetical protein C14C6.12.
 GN C14C6.12.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_Taxid=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA Wilson R.;
 RT "Genome sequence of the nematode *C. elegans*: a platform for
 RL investigating biology. The *C. elegans* Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA David M., Wohlmann P., Bauer C., Antoniou B.;
 RT "The sequence of *C. elegans* cosmid C14C6.";
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Wilson R.; (AUG-2003) to the EMBL/GenBank/DBJ databases.
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF039051; AAB94258.2; -.
 KW Hypothetical protein.
 SQ SEQUENCE 91 AA; 10407 MW; D16D2BD4011F8762 CRC64;
 QY 2 QOEAKTCHOIEYFEGDF 18
 DB 2 QDVYKICNVIEIWTGDF 18

RESULT 12
 080567 PRELIMINARY; PRT; 545 AA.
 AC 080567; Q9G5X1;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
 DE Expressed protein (VIRF-interacting protein FIP1) (Hypoetical
 DE protein) (AC2943970/F6E13.10).
 GN AT2943970 OR F6E13.10/AT2943970.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Rounsley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L.,
 RA Brandon R.C., Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D.,
 RA Somerville C.R., Venter J.C.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Lin X.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Town C.D., Kaul S.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Mayda E., Tzfira T., Citeovsky V.;
 RT "Arabidopsis thaliana VIRF-interacting protein FIP1.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,
 RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,
 RA Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,
 RA Koesema E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,
 RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
 RA Southwick A., Tracy S.E., Shinzaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RT "Full length cDNA of gene F6E13.10/AC2943970 (GI:3212854)." ;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Shinn P.,
 RA Banh J., Bowser L., Carninci P., Chung M.K., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yu G., Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Arabidopsis cDNA clones." ;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M.,
 RA Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M.,
 RA Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinzaki K.,
 RA Davis R.W., Theologis A., Ecker J.R.;
 RT "Arabidopsis cDNA clones." ;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Banh J., Bowser L.,
 RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M.,
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu S.,
 RA Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Arabidopsis ORF clones." ;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC004005; AAC23405.2; -.
 DR EMBL; AF332565; AA06847.1; -.
 DR EMBL; AY056238; AL07087.1; -.
 DR EMBL; AF375410; AAK52994.1; -.
 DR EMBL; AF367277; AAK56266.1; -.
 DR EMBL; AY129474; AAM91060.1; -.
 DR JTR; T006777; T006777.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR InterPro; IPR002344; Lupa La.
 DR InterPro; IPR006630; Lupa La_dom.
 DR Pfam; PF05383; La; 1.
 DR PRINTS; PR00302; LUPUSLA.
 KW Hypothetical protein.
 SQ SEQUENCE 545 AA; 60589 MW; E1A933261FE1ED80 CRC64;
 QY Query Match 47.1%; Score 48; DB 10; Length 545;

RA Sau S., Lee C.Y.;
 RT "Cloning of type 8 capsule genes and analysis of gene clusters for the
 RT production of different capsular polysaccharides in *Staphylococcus*
 RT aureus";
 RN J. Bacteriol. 178:2118-2126(1996).
 (2)
 RC SEQUENCE FROM N.A.
 RP STRAIN=Becker;
 RX MEDLINE=97197525;
 RA Sau S., Sun J., Lee C.Y.;
 RT "Molecular characterization and transcriptional analysis of type 8
 RT capsule genes in *Staphylococcus aureus*";
 RN J. Bacteriol. 179:1614-1621(1997).
 DR EMBL: U73374; AAA89445.1; -
 DR HSP: P27828; 1FED.
 DR GO: GO:0008761; F:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
 DR GO: GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
 DR GO: GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
 DR InterPro: IPR003331; Epimerase_2.
 DR Pfam: PF02350; Epimerase_2; 1.
 DR TIGRfams: TIGR00236; wecB; 1.
 SQ SEQUENCE 391 AA; 44247 MW; 1EBD9FA9BC76F0D CRC64;
 Query Match 51.0%; Score 52; DB 2; Length 391;
 Best Local Similarity 60.0%; Pred. No. 2.9;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 Oy 2 QOEAKECHQIEYFG 16
 Db 362 QASRRICAEIYFG 376

RESULT 9
 ID P95709 PRELIMINARY; PRT; 391 AA.
 AC P95709;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE CapS.
 GN CAPS.
 OS *Staphylococcus aureus*.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 NCBI_TaxID=1280;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Reynolds and Newman;
 RX MEDLINE=97388587; PubMed=9245821;
 RA Sau S., Bhasin N., Mann E.R., Lee J.C., Foster T.J., Lee C.Y.;
 RT "The *Staphylococcus aureus* allelic genetic loci for serotype 5 and 8
 RT capsule expression contain the type-specific genes flanked by common
 RT genes";
 RN J. Microbiol. 143:0-0(0).
 (2)
 RC SEQUENCE FROM N.A.
 RP STRAIN=Reynolds and Newman;
 RX MEDLINE=98101481; PubMed=9440531;
 RA Kiser K.B., Lee J.C.;
 RT "Staphylococcus aureus cap50 and capS genes functionally complement
 RT mutations affecting enterobacterial common-antigen biosynthesis in
 RT *Escherichia coli*";
 RN J. Bacteriol. 180:403-406(1998).
 (3)
 RC SEQUENCE FROM N.A.
 RP STRAIN=Reynolds and Newman;
 RX MEDLINE=98125727; PubMed=9466251;
 RA Bhasin N., Albus A., Michon F., Livolsi P.J., Park J.-S., Lee J.C.;
 RT "Identification of a gene essential for O-acetylation of the
 RT *Staphylococcus aureus* type 5 capsular polysaccharide";
 RN Mol. Microbiol. 27:9-21(1998).
 (4)
 RC SEQUENCE FROM N.A.
 RP STRAIN=Reynolds and Newman;

RA Bagga N., Mann E.R., Foster T.J., Lee J.C.;
 RT Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U81973; AAC46039.1; -
 DR HSP: P27828; 1FED.
 DR GO: GO:0008761; F:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
 DR GO: GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
 DR GO: GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
 DR InterPro: IPR003331; Epimerase_2.
 DR Pfam: PF02350; Epimerase_2; 1.
 DR TIGRfams: TIGR00236; wecB; 1.
 SQ SEQUENCE 391 AA; 44328 MW; B0105F690DB7CF1D CRC64;
 Query Match 51.0%; Score 52; DB 2; Length 391;
 Best Local Similarity 60.0%; Pred. No. 2.9;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 Oy 2 QOEAKECHQIEYFG 16
 Db 362 QASRRICAEIYFG 376

RESULT 10
 ID Q99X57 PRELIMINARY; PRT; 391 AA.
 AC Q99X57;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Capsular polysaccharide synthesis enzyme CapS.
 GN CAP OR SAV0164 OR SA0159.
 OS *Staphylococcus aureus* (strain Mu50 / ATCC 700699), and
 OS *Staphylococcus aureus* (strain N315).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 NCBI_TaxID=158878; 158879;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC SPECIES=S. aureus (strain Mu50), and S. aureus (strain N315);
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Ito T.,
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshiyama A.,
 RA Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hiraoka H., Kuhara S., Goto S., Yabuzaki J.,
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of methicillin-resistant *Staphylococcus*
 RT aureus";
 RN Lancet 357:1225-1240(2001).
 DR EMBL: AP003129; BAB56326.1; -
 DR EMBL: AP003129; BAB41379.1; -
 DR FIR: H89777; H89777.
 DR HSP: P27828; 1FED.
 DR GO: GO:0008761; F:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
 DR GO: GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
 DR GO: GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
 DR InterPro: IPR003331; Epimerase_2.
 DR Pfam: PF02350; Epimerase_2; 1.
 DR TIGRfams: TIGR00236; wecB; 1.
 KW Complete proteome.
 SQ SEQUENCE 391 AA; 44372 MW; DDD5FA715BCCECC CRC64;
 Query Match 51.0%; Score 52; DB 16; Length 391;
 Best Local Similarity 60.0%; Pred. No. 2.9;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 Oy 2 QOEAKECHQIEYFG 16
 Db 362 QASRRICAEIYFG 376

RESULT 11
 ID O44678 PRELIMINARY; PRT; 91 AA.

```

RESULT 5
Q8ZT10 PRELIMINARY; PRT; 401 AA.
ID Q8ZT10
AC Q8ZT10
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to sjogren syndrome antigen B (Autoantigen La).
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
NCBI_TaxID=7955;
RX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Body;
RC Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC045392; AAH45392.1; -.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR InterPro: IPR002344; Lupus_La.
DR InterPro: IPR006630; Lupus_La_dom.
DR Pfam: PF00076; rrm; 1.
DR PRINTS: PR00302; LUPUSLA.
DR SMART: SM00715; LA; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS50102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 401 AA; 46138 MW; 2EF032FDD3916291 CRC64;

Query Match
Best Local Similarity 56.9%; Score 58; DB 13; Length 401;
Pred. No. 0.31;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 EAKICHQIEYYFGD 17
Db 11 EKRVAEQIEYYFGD 24

RESULT 6
Q8T8V5 PRELIMINARY; PRT; 390 AA.
ID Q8T8V5
AC Q8T8V5
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE A122034P.
GN LA OR CG10922.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RX [1]
RN SEQUENCE FROM N.A.
RP Stapleton M., Brokstein P., Hong L., Agpayani A., Carlson J.,
RA Campe M., Chavez C., Dorett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Patagas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY05257; AAI68124.1; -.
DR FlyBase: FBgn0011638; La.
DR GO: GO:0008098; F:5S rRNA primary transcript binding; IDA.
DR GO: GO:0003723; F:RNA binding; NAS.
DR InterPro: IPR002344; Lupus_La.
DR InterPro: IPR006630; Lupus_La_dom.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF05383; La; 1.

```

```

DR Pfam: PF00076; rrm; 1.
DR PRINTS: PR00302; LUPUSLA.
DR SMART: SM00715; LA; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS50102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 390 AA; 44869 MW; 797FDE26B903C909 CRC64;

Query Match
Best Local Similarity 55.9%; Score 57; DB 5; Length 390;
Pred. No. 0.44;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 QQAKICHQIEYYFGD 17
Db 49 KQERRAIRQVEYYFGD 64

RESULT 7
Q8NYN8 PRELIMINARY; PRT; 381 AA.
ID Q8NYN8
AC Q8NYN8
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Capsular polysaccharide synthesis enzyme Cap8P.
GN CAP8P OR MW0139.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
NCBI_TaxID=196620;
RX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramoto K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA."
RL Lancet 359:1819-1827(2002).
DR EMBL: AP004822; BAB94004.1; -.
DR GO: GO:0008761; F:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
DR GO: GO:0009103; F:lipopolysaccharide biosynthesis; IEA.
DR GO: GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
DR InterPro: IPR003331; Epimerase_2.
DR Pfam: PF02350; Epimerase_2; 1.
DR TIGRFAMs: TIGR00236; wecB; 1.
KM Complete proteome.
SQ SEQUENCE 381 AA; 43106 MW; 125E4D5D1904779C CRC64;

Query Match
Best Local Similarity 51.0%; Score 52; DB 16; Length 381;
Pred. No. 2.9;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 QQAKICHQIEYYFG 16
Db 352 QASRRICEAIEYYFG 366

RESULT 8
P72382 PRELIMINARY; PRT; 391 AA.
ID P72382
AC P72382
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CAP8P.
GN CAP8P.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
NCBI_TaxID=1280;
RX [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Becker;
MEDLINE=96178981; PubMed=8606192;

```


DR GO:0005634; C:nucleus; IDA.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rtm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 SQ SEQUENCE 381 AA; 43891 MW; 2E2DEFF1452C0F0E9 CRC64;

Query Match 85.3%; Score 87; DB 11; Length 381;
 Best Local Similarity 100.0%; Pred. No. 5.2e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
 DB 14 EAKICHOIEYFGDF 28

RESULT 2

Q8BTU4 PRELIMINARY; PRT; 415 AA.
 AC Q8BTU4;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Sjogren syndrome antigen B.
 GN SSB.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MOD; TISSUE=Thymus;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The PANThOM Consortium;
 RT The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 DR EMBL; AK086677; BAC040498.1; -
 DR MGD; MGI:98423; Ssb.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rtm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 SQ SEQUENCE 415 AA; 47657 MW; A7545CT686AC8363 CRC64;

Query Match 85.3%; Score 87; DB 11; Length 415;
 Best Local Similarity 100.0%; Pred. No. 5.7e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
 DB 14 EAKICHOIEYFGDF 28

RESULT 3

Q7ZTK2 PRELIMINARY; PRT; 427 AA.
 AC Q7ZTK2;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Similar to lupus LA protein homolog B.

OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=embryo;
 RA Klein S., Strassberg R.;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC046654; AA046654.1; -
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rtm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 2.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 SQ SEQUENCE 427 AA; 48996 MW; 1E7CD82D8AB9C69A CRC64;

Query Match 64.7%; Score 66; DB 13; Length 427;
 Best Local Similarity 78.6%; Pred. No. 0.016;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGD 17
 DB 13 DTKICEQIEYFGD 26

RESULT 4
 Q8QCHIS PRELIMINARY; PRT; 206 AA.
 ID Q8QCHIS;
 AC Q8QCHIS;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Untranslated region binding protein.
 GN UBP.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA L'Ecuier T.J., Fang H.-L.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF467897; AAL76269.1; -
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR Pfam; PF00076; rtm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 SQ SEQUENCE 206 AA; 23992 MW; 965B62F7DFB90E9 CRC64;

Query Match 63.7%; Score 65; DB 13; Length 206;
 Best Local Similarity 78.6%; Pred. No. 0.011;
 Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGD 17
 DB 14 ESKICQIEYFGN 27

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:29:01 ; Search time 41.229 Seconds
(without alignments)
137.751 Million cell updates/sec

Title: US-09-836-073-4

Perfect score: 102
Sequence: 1 QOQEKICHIQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTRMBL_25:*
1: sp_bacteria:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	85.3	381	11	Q9CYB9
2	87	85.3	415	11	Q8BRTU4
3	66	64.7	427	13	Q7ZTK2
4	65	63.7	206	13	Q8OH15
5	58	56.9	401	13	Q7ZTK10
6	57	55.9	390	5	Q8T8V5
7	52	51.0	381	16	Q8NYS8
8	52	51.0	391	2	P95709
9	52	51.0	391	16	Q9X57
10	52	51.0	391	16	Q9X57
11	50	49.0	5	5	Q4678
12	48	47.1	545	10	Q80567
13	48	47.1	756	16	Q8ABY5
14	48	47.1	2349	5	Q81455
15	47.5	46.6	329	16	Q87R57
16	47	46.1	396	5	Q01806

17	47	46.1	411	10	Q9FL36
18	47	46.1	422	10	Q94A38
19	47	46.1	472	16	Q8D9F3
20	47	46.1	547	10	Q7XVC6
21	46	45.1	478	16	Q8TGS8
22	46	45.1	642	16	Q9KEU7
23	45	44.1	352	16	Q8YMW5
24	45	44.1	481	10	Q8LMP9
25	45	44.1	568	16	Q9KXL7
26	45	44.1	749	2	Q84ID5
27	45	44.1	1330	5	Q9GQ43
28	45	44.1	1858	5	Q8MSU5
29	45	44.1	2145	5	Q9W003
30	44.5	43.6	114	16	Q7UT07
31	44.5	43.6	928	10	Q9LJ02
32	44	43.1	96	10	Q9AUG1
33	44	43.1	329	5	Q8T907
34	44	43.1	329	5	Q9VAX9
35	44	43.1	383	10	Q7YIE1
36	44	43.1	472	5	Q95PU7
37	44	43.1	1401	5	Q81SD4
38	43.5	42.6	788	17	Q58603
39	43.5	42.6	1037	16	Q8XE2
40	43.5	42.6	1037	16	Q8TF84
41	43.5	42.6	1037	16	Q83QMS
42	43	42.2	119	10	Q9ZP89
43	43	42.2	240	16	Q8RE78
44	43	42.2	255	5	Q9NGG5
45	43	42.2	255	5	Q9NEL7

ALIGNMENTS

RESULT 1
ID Q9CYB9 PRELIMINARY; PRT; 381 AA.
AC Q9CYB9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sjogren syndrome antigen B.
GN SSB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh Y.,
RA Atakawa T., Hara A., Fukunishi Y., Konno H., Adachi Y., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
RA Kiehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiraldi L.M., Staudli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,
RA Wuzhuh-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohatsu S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:885-690(2001).
DR EMBL; AK017822; BAB30957.1; -;
DR MGD; MGI:98423; SSB.

RX MEDLINE=22388257; PubMed=12477932;

RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klansner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loguillano N.A., Peters G.J., Abramson R.D., Mullish S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kerteman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,

RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length

RT human and mouse cDNA sequences.";

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -1- SIMILARITY: Contains 1 RING-type zinc finger.

CC -1- SIMILARITY: Contains 1 B box-type zinc finger.

CC -1- SIMILARITY: Contains 1 SPRY domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U09825; AAA93131.1; -

DR EMBL; AP000517; BAB63330.1; -

DR EMBL; BC024039; AAH32297.1; -

DR EMBL; BC032297; AAH32297.1; -

DR Genew; HGNC:12962; TRIM26.

DR MIM; 600830; -

DR GO; GO:0003677; F:DNA binding; TAS.

DR GO; GO:0005515; F:protein binding; TAS.

DR InterPro; IPR001870; B302.

DR InterPro; IPR006574; PRY.

DR InterPro; IPR003877; SPRY_receptor.

DR InterPro; IPR000315; Znf_Box.

DR Pfam; PF00622; SPRY; 1.

DR Pfam; PF00643; Zf-B_box; 1.

DR PRINTS; PR01406; BBOXZNFINGER.

DR SMART; SM00336; BBOX; 1.

DR SMART; SM00589; PRY; 1.

DR SMART; SM00184; RING; 1.

DR SMART; SM00449; SPRY; 1.

DR PROSITE; PSS0119; ZF_BOX; 1.

DR PROSITE; PSS00518; ZF_RING; 1.

DR PROSITE; PSS0089; ZF_RING; 1.

KW zinc-finger; Coiled coil.

FT ZN_FING 16 57 RING-TYPE.

FT ZN_FING 97 138 B_BOX-TYPE.

FT DOMAIN 188 227 COILED COIL (POTENTIAL).

FT DOMAIN 365 539 SPRY.

FT DOMAIN 388 400 POLY-GLU.

FT DOMAIN 422 432 POLY-GLU.

SQ SEQUENCE 539 AA; 62165 MW; 842A71C41F2E2348 CRC64;

Query Match 40.2%; Score 41; DB 1; Length 539;

Best Local Similarity 38.1%; Pred. No. 34;

Matches 8; Conservative 5; Mismatches 4; Indels 4; Gaps 1;

OY 1 000EAKIC---HQIEYFGD 17

DB 95 EOODATLCEHREKATHYCED 115

Search completed: September 10, 2004, 17:53:04
Job time : 8.24022 secs

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NUB36;
 RX MEDLINE=91310597; PubMed=1856180;
 RA Wu L., Welker N.E.;
 RT "Cloning and characterization of a glutamine transport operon of
 RT Bacillus stearotherophilus NUB36: effect of temperature on
 RT regulation of transcription.";
 RL J. Bacteriol. 173:4877-4888 (1991).
 CC -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
 CC FOR GLUTAMINE. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO THE
 CC TRANSPORT SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Membrane-associated (Potential).
 CC -1- INDUCTION: By lack of glutamine.
 CC -1- SIMILARITY: Belongs to the ABC transporter family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M61017; AAA22483.1; -.
 DR PIR; A42478; A42478.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003439; ABC_Transporter.
 DR Pfam; PF00005; ABC_tran; 1.
 DR ProDom; PD000006; ABC_transporter; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
 DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
 DR Amino-acid transporter; Transporter; Membrane; ATP-binding.
 FT NP BIND 34 41 ATP (By SIMILARITY).
 SQ SEQUENCE 242 AA; 27436 MW; 102B1C5B332F31C8 CRC64;
 Query Match 41.2%; Score 42; DB 1; Length 242;
 Best Local Similarity 60.0%; Pred. No. 9.8;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 9 HQLEYYFGDF 18
 DB 5 HQVKNKYGDF 14

RT protein homologous to the human placental isoform in U1210 murine
 RT leukemia cell lines with a defective reduced folate carrier.";
 RL J. Biol. Chem. 269:4267-4272 (1994).
 CC -1- FUNCTION: Binds to folate and reduced folic acid derivatives and
 CC mediates delivery of 5-methyltetrahydrofolate to the interior of
 CC cells.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
 CC similarity).
 CC -1- SIMILARITY: Belongs to the folate receptor family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M64817; AAA37599.1; -.
 DR PIR; L25338; AAA37594.1; -.
 DR PIR; B40969; B40969.
 DR MGI; MGI:95569; Folr2.
 DR InterPro; IPR004269; Folate_recep.
 DR Pfam; PF03024; Folate_rec; 1.
 DR Receptor; Glycoprotein; Signal; Placenta; Folate-binding; Membrane;
 KW GPI-anchor; Multigene family; Lipoprotein.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 227 FOLATE RECEPTOR BETA.
 FT PROPEP 228 251 REMOVED IN MATURE FORM (POTENTIAL).
 FT LIPID 227 227 GPI-anchor attached serine (POTENTIAL).
 FT CARBOHYD 62 62 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 193 193 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 251 AA; 28821 MW; 8404EACB1BFECC7 CRC64;
 Query Match 40.2%; Score 41; DB 1; Length 251;
 Best Local Similarity 75.0%; Pred. No. 15;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 8 CHQIEYYF 15
 DB 167 CHTFEYF 174

RESULT 15
 FOL2_MOUSE STANDARD; PRT; 539 AA.
 ID_2173_HUMAN
 AC Q12899;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Zinc finger protein 173 (Tripartite motif-containing protein 26) (Acid
 DE finger protein) (AFP).
 GN TRIM26 OR ZNF173.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96079113; PubMed=8530076;
 RA Chu T.W., Caposela A., Coleman R., Goel V.L., Nallur G., Gruen J.R.;
 RT "Cloning of a new 'finger' protein gene (ZNF173) within the class I
 RT region of the human MHC.";
 RL Genomics 29:229-239 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Shima S., Tamaya G., Oka A., Inoko H.;
 RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain, and skin;

```

CC -1- FUNCTION: PARTICIPATES IN THE EFFLUX OF AMINOGLYCOSIDES. CONFERES
CC RESISTANCE TO A VARIETY OF THESE SUBSTANCES.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ACRB/ACRD/ACRF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U12598; AAA20584.1; -
DR EMBL; U10436; AAA74741.1; -
DR EMBL; AE000334; AAC75523.1; -
DR EMBL; D90875; BAA16344.1; -
DR EMBL; D90876; BAA16348.1; -
DR EMBL; X57403; CAA40663.1; -
DR PIR; E65022; E65022.
DR Ecocore; EGI0014; acrd.
DR InterPro; IPR001036; Acrflvin_res.
DR Pfam; PF00873; ACR_tran.1.
DR PRINTS; PR00702; ACRIFLAVINRP.
DR TIGRFAMs; TIGR00915; 2A0602; 1.
DR Transmembrane; Inner membrane; Transport; Complete proteome.
FT DOMAIN 1 9
FT TRANSMEM 29 28
FT TRANSMEM 340 339
FT TRANSMEM 360 359
FT TRANSMEM 366 365
FT TRANSMEM 386 385
FT TRANSMEM 391 391
FT TRANSMEM 392 413
FT TRANSMEM 414 441
FT TRANSMEM 442 460
FT TRANSMEM 461 473
FT TRANSMEM 474 496
FT TRANSMEM 497 537
FT TRANSMEM 538 556
FT TRANSMEM 557 870
FT TRANSMEM 871 880
FT TRANSMEM 891 896
FT TRANSMEM 897 916
FT TRANSMEM 917 922
FT TRANSMEM 923 944
FT TRANSMEM 945 971
FT TRANSMEM 972 990
FT TRANSMEM 991 1003
FT TRANSMEM 1004 1026
FT TRANSMEM 1027 1037
FT TRANSMEM 1037 303
FT TRANSMEM 303 372
FT TRANSMEM 372 385
FT TRANSMEM 385 461
FT TRANSMEM 461 665
FT TRANSMEM 665 763
FT TRANSMEM 763 775
FT TRANSMEM 775 778
FT TRANSMEM 778 778
SQ SEQUENCE 1037 AA; 113047 MW; 961611E1D24FD4B5 CRC64;

Query Match 42.6%; Score 43.5; DB 1; Length 1037;
Best Local Similarity 62.5%; Pred. No. 26;
Matches 10; Conservative 1; Mismatches 4; Indels 1; Gaps 1;
QY 1 QOQKAKICHOIE-YTF 15
DB 583 QOQTLKVEQIEKYTF 598

```

```

MEI9 DROME
ID MEI9 DROME STANDARD; PRT; 926 AA.
AC Q24087;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Mei-9 protein (Meiotic-9 protein).
GN MEI-9.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriidae; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RX MEDLINE=96109608; PubMed=8647398;
RA Sekelsky J.J., McKim K.S., Chin G.M., Hawley R.S.;
RT "The Drosophila meiotic recombination gene mei-9 encodes a homologue
RT of the yeast excision repair protein Rad1."
RL Genetics 141:619-627(1995).
CC -1- FUNCTION: Implicated in recombination events during meiosis,
CC mostly in meiotic exchange. May directly resolve Holliday
CC junctions within recombination intermediates leading to DNA
CC exchange. Also required for the repair of mismatches within
CC meiotic heteroduplex DNA and for nucleotide excision repair.
CC -1- SUBUNIT: Heterodimer (by similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: Belongs to the XPF family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U27181; AAC46917.1; -
DR PIR; S58936; S58936.
DR FlyBase; FBgn002707; mei-9.
DR GO; GO:0007059; P:chromosome segregation; IMP.
DR GO; GO:0007131; P:meiotic recombination; IMP.
DR GO; GO:0006298; P:mismatch repair; IMP.
DR GO; GO:0006289; P:nucleotide-excision repair; IMP.
DR InterPro; IPR006166; ERCC4.
DR InterPro; IPR006167; Rad_1.
DR Pfam; PF02732; ERCC4; 1.
DR TIGRFAMs; TIGR00596; rad1; 1.
KW Meiosis; DNA repair; DNA-binding; Nuclear protein; Hydrolase;
KW Nuclease; Endonuclease.
SQ SEQUENCE 926 AA; 105685 MW; 342C2C91BB3AFD91 CRC64;

Query Match 42.2%; Score 43; DB 1; Length 926;
Best Local Similarity 50.0%; Pred. No. 28;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 3 QEAKICHOIEY 14
DB 371 QDARTCHOLKQY 382

RESULT 13
ID GINO BACST STANDARD; PRT; 242 AA.
AC P27675;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutamine transport ATP-binding protein glnO.
GN GINO.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
OX NCBI_Taxid=1422;

```

DR PIR; T43542; T43542.
 DR Genedb_Spombe: SPAC57A10.10c; -
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dcm.
 DR InterPro; IPR000504; RNA_rec_moc.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; La; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS0102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
 DR RNA-binding; Nuclear protein.
 DR RNA-binding; Nuclear protein.
 FT DOMAIN 154 235 M -> I (IN REF. 1 AND 2).
 FT CONFICT 188 188
 SQ SEQUENCE 298 AA; 34616 MW; 6486AB99940B87F4 CRC64;
 Query Match 45.1%; Score 46; DB 1; Length 298;
 Best Local Similarity 50.0%; Pred. No. 2.6;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 4 EAKICHOIEYFGD 17
 DB 64 EAEVLKQVERFYFD 77
 RESULT 10
 ID TCMO_PETCR STANDARD; PRT; 506 AA.
 AC 04303;
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 42, Last annotation update)
 DE 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Trans-cinnamate 4-monooxygenase (EC 1.14.13.11) (Cinnamic acid
 4-hydroxylase) (C4H) (P450C4H) (Cytochrome P450 73).
 GN CYP7A10 OR CYP73.
 OS Petroselinum crispum (Parsley) (Petroselinum hortense).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Campanulids; Apiales; Apiaceae; Apioidae; apioid superclade;
 OC Apium clade; Petroselinum.
 OC NCBI_TaxID=4043;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95320184; PubMed=7597051;
 RA Logemann E., Parniske M., Hahlbrock K.;
 RT "Modes of expression and common structural features of the complete
 phenylalanine ammonia-lyase gene family in parsley.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:5905-5909(1995).
 CC -1- FUNCTION: Controls carbon flux to pigments essential for
 pollination or UV protection, to numerous phycoalexins synthesized
 by plants when challenged by pathogens, and to lignins.
 CC -1- CATALYTIC ACTIVITY: Trans-cinnamate + NADPH + O(2) = 4-
 hydroxycinnamate + NADP(+) + H(2)O.
 CC -1- PATHWAY: Lignin biosynthesis.
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by, and for commercial
 entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; L38898; AAC41660.1; -
 DR PIR; T14907; T14907.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 RN OXidoreductase; Monooxygenase; Heme; NADP.

FT METAL 448 448 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 506 AA; 58047 MW; 32F0EE959D69CCF CRC64;
 Query Match 44.1%; Score 45; DB 1; Length 506;
 Best Local Similarity 37.5%; Pred. No. 6.8;
 Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 3 OEAKICHOIEYFGDF 18
 DB 213 ERSRLAQSPFHYHGF 228
 RESULT 11
 ID ACRD_ECOLI STANDARD; PRT; 1037 AA.
 AC P24177; P76971; P77178; 046715;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable aminoglycoside efflux pump (Acriflavine resistance protein
 D).
 GN ACRD OR B2470.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OC NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RA Niles M.L., Bertrand K.P.;
 RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ma D., Cook D.N., Alberti M., Nikaido H., Hearst J.E.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Colado-Vides J., Glaser J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1233-1238(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97349980; PubMed=9205837;
 RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,
 RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
 RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
 RA Osahima T., Oyama S., Saito N., Sempel G., Satoh Y., Sivasubraman S.,
 RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
 RA Yamagata S., Horiiuchi T.;
 RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
 K-12 genome corresponding to 50.0-68.8 min on the linkage map and
 analysis of its sequence features.";
 RL DNA Res. 4:91-113(1997).
 RN [5]
 RP SEQUENCE OF 998-1037 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=92355499; PubMed=1644752;
 RA Bouvier J., Richard C., Higgins W., Bogler O., Stragier S.;
 RT "Cloning, characterization, and expression of the *dapB* gene of
 Escherichia coli.";
 RL J. Bacteriol. 174:5265-5271(1992).
 RN [6]
 RP FUNCTION.
 RX MEDLINE=20158894; PubMed=10692383;
 RA Rosenberg E.Y., Ma D., Nikaido H.;
 RT "AcrD of Escherichia coli is an aminoglycoside efflux pump.";
 RL J. Bacteriol. 182:1754-1756(2000).

FT CONFLICT 182 183 KH -> NS (IN REF. 1).
 FT CONFLICT 283 283 A -> R (IN REF. 1).
 FT CONFLICT 329 329 K -> N (IN REF. 1).
 SQ SEQUENCE 390 AA; 44884 MW; A809288B90446A5 CRC64;
 Query Match 55.9%; Score 57; DB 1; Length 390;
 Best Local Similarity 62.5%; Pred. No. 0.052;
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 QOEAKIQIEYFGD 17
 : |||||
 DB 49 KOERAIRQVEYFGD 64

RESULT 8
 LA AEDAL STANDARD; PRT; 383 AA.
 AC Q26457;
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE A protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 OS Aedes albopictus (Forest day mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.
 OX NCBI_TaxID=7160;
 RN NCBI
 RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=96135233; PubMed=8551578;
 RA Pardigon N., Straus J.H.;
 RT "Mosquito homolog of the La autoantigen binds to Sindbis virus RNA.";
 RL J. Virol. 70:1173-1181(1996).
 CC -1- FUNCTION: May be involved in transcription termination by RNA
 minus strand of Sindbis virus RNA. Binds to the 3' end of the
 CC Sindbis virus RNA replication.
 CC -1- SUBCELLULAR LOCATION: Nuclear. Primarily nuclear, but significant
 CC amounts are present in the cytoplasm.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; S80954; AAB35931.1; -;
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PR00076; rrm, 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
 KM RNA-binding; Nuclear Protein; DNA-binding.
 FT DOMAIN 141 228
 SQ SEQUENCE 383 AA; 44430 MW; 4E5CC8F21C40F452 CRC64;

Query Match 50.0%; Score 51; DB 1; Length 383;
 Best Local Similarity 64.3%; Pred. No. 0.51;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 EAKICQIEYFGD 17
 : |||||
 DB 44 EASTIRQLEYYFGD 57

RESULT 9
 LAH1 SCHPO STANDARD; PRT; 298 AA.
 AC P87058; Q10458;
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 GN SLA1 OR SPAC57A10.10C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN NCBI
 RN SEQUENCE FROM N.A.
 RX MEDLINE=98067398; PubMed=9404894;
 RA Van Horn D.J., Yoo C.J., Xue D., Shi H., Wollin S.L.;
 RT "The La protein in Schizosaccharomyces pombe: a conserved yet
 RT dispensable phosphoprotein that functions in tRNA maturation.";
 RL RNA 3:1434-1443(1997).
 RN (2)
 RN SEQUENCE FROM N.A.
 RA Utsunomi R.R.U.;
 RA Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RN SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Church C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leathers S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkhardt G., Aert R., Robben J., Grymporter B.,
 RA Welcens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Meestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Egger P., Zimmermann W., Wedler H., Wandut R., Fumelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucban M., Rochet M., Galliard C., Tallada V.A., Garçon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armitrong J., Forburg S.L.,
 RA Cervetti L., Lowe T., McCombie W.R., Paulsen I., Pechankin J.,
 RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: Binds to the precursors of polymerase III RNAs.
 CC Functions in tRNA maturation.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF022949; AAB82145.1; -;
 DR EMBL; AB011371; BAA24981.1; -;
 DR EMBL; Z94864; CAB08173.1; -;
 DR PIR; T38937; T38937.

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X68817; CAA48715.1; -

DR PIR; S33818; S33818.

DR InterPro; IPR002344; Lupus_La.

DR InterPro; IPR006630; Lupus_La_dom.

DR InterPro; IPR000504; RNA_rec_mot.

DR Pfam; PF05383; Rrm; 1.

DR Pfam; PF00076; Rrm; 1.

DR PRINTS; PR00302; LUPUSIA.

DR SMART; SMO0715; LA; 1.

DR SMART; SMO0360; RRM; 1.

DR PROSITE; PS50102; RRM; 1.

DR PROSITE; PS00030; RRM_RNP_1; 1.

DR RNA-binding; Nuclear Protein; Phosphorylation.

FW DOMAIN 111 203 RNA-BINDING (RRM).

FT DOMAIN 196 212 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

FT DOMAIN 316 332 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

SEQUENCE 428 AA; 48864 MW; AEB3A38B7D2E3EC3 CRC64;

Query Match 64.7%; Score 66; DB 1; Length 428;
Best Local Similarity 78.6%; Pred. No. 0.0018;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGD 17
: ||| |||||
Db 14 DTKICEQIEYFGD 27

RESULT 7

LA DROME STANDARD; PRT; 390 AA.

AC P40796; Q24375; Q9VIN2;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
GN LA OR CG10922.

OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S; TISSUE=Ovary;
RX MEDLINE=94309632; PubMed=8035794;
RA Bai C., Li Z., Tolia P.P.;
RT "Developmental characterization of a Drosophila RNA-binding protein homologous to the human systemic lupus erythematosus-associated La/SS-B autoantigen."
RT Mol. Cell. Biol. 14:5123-5129(1994).
RL [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=94309661; PubMed=8035818;
RA Yoo C.J., Wolin S.L.;
RT "A proteins from Drosophila melanogaster and Saccharomyces cerevisiae: a yeast homolog of the La autoantigen is dispensable for growth."
RT Mol. Cell. Biol. 14:5412-5424(1994).
RL [3]
RN SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.P.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brudon R.C., Rogers J.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abtil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Balliw R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhattacharya D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Bottier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodika C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lascko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Mostreli A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murry D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: May be involved in transcription termination by RNA polymerase III. Binds RNA and DNA. Binds to precursors of RNA development. III transcripts. May play a specialized role during fly development.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- DEVELOPMENTAL STAGE: Expressed throughout embryonic, larval, pupal, and adult development. Expression throughout the embryo is followed by a restricted pattern of mesodermal expression that is later confined to the visceral mesoderm, gonads, gut, and salivary glands.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U07652; AAA20518.1; -

DR EMBL; L32988; AAA21776.1; -

DR EMBL; AEO03666; AAF53885.1; -

DR PIR; A53773; A53773.

DR PIR; A53781; A53781.

DR FLYBase; FBgn0011638; La.

DR GO; GO:0008098; F:5S rRNA primary transcript binding; IDA.

DR GO; GO:0003723; F:5S rRNA binding; NAS.

DR InterPro; IPR002344; Lupus_La.

DR InterPro; IPR006630; Lupus_La_dom.

DR InterPro; IPR000504; RNA_rec_mot.

DR Pfam; PF05383; Rrm; 1.

DR Pfam; PF00076; Rrm; 1.

DR PRINTS; PR00302; LUPUSIA.

DR SMART; SMO0715; LA; 1.

DR SMART; SMO0360; RRM; 1.

DR PROSITE; PS50102; RRM; 1.

DR PROSITE; PS00030; RRM_RNP_1; 1.

DR RNA-binding; Nuclear Protein; DNA-binding.

FW DOMAIN 149 234 RNA-BINDING (RRM).

FT DOMAIN 169 169 A -> T (IN REF. 1).

FT CONFLICT 169 169


```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X67859; CAA8043.1; -
CC PIR; JCI494; JCI494.
CC InterPro; IPR002344; Lupus_La.
CC InterPro; IPR006630; Lupus_La.dom.
CC InterPro; IPR000504; RNA_rec_mot.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF05383; La; 1.
CC Pfam; PF00076; xrm; 1.
CC SMART; SM00715; LA; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00715; LA; 1.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS50102; RRM; 1.
CC PROSITE; PS00030; RRM_RNP_1; 1.
CC RNA-binding; Nuclear protein; Phosphorylation.
CC FT DOMAIN 111 187 RNA-BINDING (RRM).
CC SQ SEQUENCE 415 AA; 47777 MW; 033FD9CC1E475F98 CRC64;

Query Match 85.3%; Score 87; DB 1; Length 415;
Best Local Similarity 100.0%; Pred. No. 5.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 EAKICHQIEYFGDF 18
DB 14 EAKICHQIEYFGDF 28

RESULT 5
LAB_XENLA STANDARD; PRT; 427 AA.
ID LAB_XENLA
AC P28049;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Lupus La protein homolog B (La ribonucleoprotein B) (La autoantigen
DE homolog B).
GN LAB1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OC NCB1_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RX MEDLINE=93287095; PubMed=8510143;
RA Scherly D., Stutz F., Lin-Marg N., Clarkson S.G.;
RT "La proteins from Xenopus laevis. cDNA cloning and developmental
RT expression.";
RL J. Mol. Biol. 231:196-204(1993).
CC -1- FUNCTION: La protein plays a role in the transcription of RNA
CC polymerase III. It is most probably a transcription termination
CC factor. Binds to the 3' terminus of virtually all nascent
CC polymerase III transcripts (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- DEVELOPMENTAL STAGE: Barely detectable in stage I/II oocytes,
CC accumulate in stage III/IV oocytes, then exhibit a roughly
CC constant steady state level in mature oocytes, eggs, and early
CC embryos.
CC -1- PTM: Phosphorylated (Probable).
CC -1- MISCELLANEOUS: There are two forms of La, Laa and Lab, in xenopus.
CC -1- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X68818; CAA48716.1; -
CC PIR; S33817; S33817.
CC InterPro; IPR002344; Lupus_La.
CC InterPro; IPR006630; Lupus_La.dom.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF05383; La; 1.
CC Pfam; PF00076; xrm; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00715; LA; 1.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS50102; RRM; 1.
CC PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
CC RNA-binding; Nuclear protein; Phosphorylation.
CC FT DOMAIN 110 202 RNA-BINDING (RRM).
CC FT DOMAIN 315 331 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC SQ SEQUENCE 427 AA; 48995 MW; 45F3146F8934A355 CRC64;

Query Match 64.7%; Score 66; DB 1; Length 427;
Best Local Similarity 78.6%; Pred. No. 0.0018;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 EAKICHQIEYFGDF 17
DB 13 DTKICQIEYFGDF 26

RESULT 6
LAA_XENLA STANDARD; PRT; 428 AA.
ID LAA_XENLA
AC P28048;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Lupus La protein homolog A (La ribonucleoprotein A) (La autoantigen
DE homolog A).
GN LAA1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OC NCB1_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RX MEDLINE=93287095; PubMed=8510143;
RA Scherly D., Stutz F., Lin-Marg N., Clarkson S.G.;
RT "La proteins from Xenopus laevis. cDNA cloning and developmental
RT expression.";
RL J. Mol. Biol. 231:196-204(1993).
CC -1- FUNCTION: La protein plays a role in the transcription of RNA
CC polymerase III. It is most probably a transcription termination
CC factor. Binds to the 3' terminus of virtually all nascent
CC polymerase III transcripts (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- DEVELOPMENTAL STAGE: Barely detectable in stage I/II oocytes,
CC accumulate in stage III/IV oocytes, then exhibit a roughly
CC constant steady state level in mature oocytes, eggs, and early
CC embryos.
CC -1- PTM: Phosphorylated (Probable).
CC -1- MISCELLANEOUS: There are two forms of La, Laa and Lab, in xenopus.
CC -1- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

DR PROSITE: PS00030; RRM_RNP_1; 1.
 KW Systemic lupus erythematosus; RNA-binding; Phosphorylation;
 KM Nuclear protein.
 FT DOMAIN 111 187 RNA-BINDING (RRM).
 FT MOD RES 366 366 PHOSPHORYLATION (BY CK2).
 SQ SEQUENCE 408 AA; 46837 MW; EC153C15F9187FC4 CRC64;
 Query Match 85.3%; Score 87; DB 1; Length 408;
 Best Local Similarity 100.0%; Pred. No. 5.4e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
 DB 14 EAKICHOIEYFGDF 28

RESULT 3
 LA MOUSE STANDARD; PRT; 415 AA.
 ID LA MOUSE
 AC P32067;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus la protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 GN SSB OR SS-B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RX SEQUENCE FROM N.A.
 RP MEDLINE=93203630; PubMed=8454877;
 RA Topfer F., Gordon T., McCluskey J.;
 RT "Characterization of the mouse autoantigen La (SS-B). Identification of conserved RNA-binding motifs, a putative ATP binding site and reactivity of recombinant protein with poly(U) and human autoantibodies.";
 RT J. Immunol. 150:3091-3100(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary gland;
 RX MEDLINE=22388257; PubMed=12479932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stacheron M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda N.T., Toshitoki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mulhally S.J.,
 RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hilyk S.W.,
 RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Heiton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE OF 1-11 FROM N.A.
 RA Grodzil D., Bachmann M.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: La protein plays a role in the transcription termination of polymerase III. It is most probably a transcription termination factor. Binds to the 3' terminus of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S, and 7-2 RNAs.

CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: Phosphorylated (By similarity).
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 CC EMBL; L00993; AAA39415.1; -
 CC EMBL; BC003820; AAH03820.1; -
 CC EMBL; Y07951; CAA69249.1; -
 CC MGI; MGI:98423; Ssb.
 CC GO; GO:0005634; C:nucleus; IDA.
 CC InterPro; IPR002344; Lupus_La.
 CC InterPro; IPR006630; Lupus_La_dom.
 CC InterPro; IPR000504; RNA_rec_mot.
 CC Pfam; PF05383; La; 1.
 CC Pfam; PF0076; Rrm; 1.
 CC PRINTS; PR00302; IUPUSLA.
 CC SMART; SM00715; LA; 1.
 CC SMART; SM00360; RRM; 1.
 CC DR PROSITE; PSS0102; RRM; 1.
 CC DR PROSITE; PSS0030; RRM_RNP_1; 1.
 KW RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 111 187 RNA-BINDING (RRM).
 PT SEQUENCE 415 AA; 47756 MW; 2D75197692DFC933 CRC64;
 SQ SEQUENCE

Query Match 85.3%; Score 87; DB 1; Length 415;
 Best Local Similarity 100.0%; Pred. No. 5.5e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
 DB 14 EAKICHOIEYFGDF 28

RESULT 4
 LA RAT STANDARD; PRT; 415 AA.
 ID LA RAT
 AC P38656;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lupus la protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 GN SSB OR SS-B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93246255; PubMed=7916708;
 RA Senses I., Troester H., Bartesch H., Schwemmler M., Igloi G.L.,
 RA Bachmann M.;
 RT "Isolation of rat cDNA clones coding for the autoantigen SS-B/La: detection of species-specific variations.";
 RT Gene 126:265-268(1993).
 RL Gene 126:265-268(1993).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' terminus of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S, and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: Phosphorylated (By similarity).
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.

FT DOMAIN 111 187 RNA-BINDING (RM).
 SQ SEQUENCE 404 AA; 46534 MW; 4EE30B5C262AD6A1 CRC64;
 Query Match 85.3%; Score 87; DB 1; Length 404;
 Best Local Similarity 100.0%; Pred. No. 5.3e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHQIERYFGDF 18
 DB 14 EAKICHQIERYFGDF 28

RESULT 2
 LA HUMAN STANDARD; PRT; 408 AA.
 ID P05455;
 AC 01-NOV-1988 (Rel. 09, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein (Sjogren syndrome type B antigen) (SS-B) (La ribonucleoprotein) (La autoantigen).
 GN SSB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Eureleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NC NCB1_TaxID=9606;
 OK NCB1_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89202037; PubMed=2468131;
 RT Chan E.K.L., Sullivan K.F., Tan E.M.;
 RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus Nucleic Acids Res. 17:2233-2244(1989)."
 RL (2)
 RN SEQUENCE FROM N.A.
 RX MEDLINE=89053970; PubMed=3192525;
 RT Chambers J.C., Kenan D., Martin B.J., Keene J.D.;
 RT "Genomic structure and amino acid sequence domains of the human La autoantigen.";
 RL J. Biol. Chem. 263:18043-18051(1988).
 RN (3)
 RN SEQUENCE FROM N.A.
 RC TISSUE=Placenta, and Skeletal muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallary S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 RA Villalón D.K., Muray D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalka U., Smaluk D.E.,
 RA Schermer A., Schein J.E., Jones S.J.M., Merra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN (4)
 RN SEQUENCE OF 54-408 FROM N.A.
 RP MEDLINE=88199081; PubMed=2452201;
 RX Scurgess A.D., Peterson M.G., McNeillage L.J., Whittingham S.,
 RA Coppel R.S.;
 RT "Characteristics and epitope mapping of a cloned human autoantigen La.";
 RL J. Immunol. 140:3212-3218(1988).
 RN (5)

RP SEQUENCE OF 54-97 FROM N.A.
 RX MEDLINE=85166283; PubMed=3856888;
 RA Chambers J.C., Keene J.D.;
 RT "Isolation and analysis of cDNA clones expressing human lupus La antigen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2115-2119(1985).
 RN (6)
 RP FUNCTION.
 RX MEDLINE=89251617; PubMed=2470590;
 RA Gottlieb E., Steltz J.A.;
 RT "Function of the mammalian La protein: evidence for its action in transcription termination by RNA polymerase III.";
 RL EMBO J. 8:851-861(1989).
 RN (7)
 RP PHOSPHORYLATION.
 RX MEDLINE=92707017; PubMed=9054510;
 RA Fan H., Sakulich A.L., Goodier J.L., Zhang X., Qin J., Marrate R.J.;
 RT "Phosphorylation of the human La antigen on serine 366 can regulate recycling of RNA polymerase III transcription complexes.";
 RL Cell 88:707-715(1997).
 RN (8)
 RP INTERACTION WITH DDX15.
 RX MEDLINE=22346609; PubMed=12458796;
 RA Fournaux M.A., Kolkman M.J.M., Van der Heijden A., De Jong A.S.,
 RA Van Venrooij W.J., Pruijn G.J.M.;
 RT "The human La (SS-B) autoantigen interacts with DDX15/hPrp43, a putative DEAD-box RNA helicase.";
 RL RNA 8:1428-1443(2002).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' termini of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S, and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE C-TERMINAL PART OF THE PROTEIN.
 CC -1- PTM: The N-terminus is blocked.
 CC -1- DISEASE: Sera from patients with systemic lupus erythematosus often contain antibodies that react with the normal cellular La protein as if this antigen was foreign.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X13697; CAA31985.1; -;
 DR EMBL: J04205; AAA51885.1; -;
 DR EMBL: BC001289; AAH01289.1; -;
 DR EMBL: BC020818; AAH20818.1; -;
 DR PIR: A31888; A31888.
 DR Genew: HGNC:11316; SSB.
 DR MIM: 109090; -;
 DR GO: GO:0030529; C:ribonucleoprotein complex; TAS.
 DR GO: GO:0003729; F:RNA binding; TAS.
 DR GO: GO:0000049; F:RNA binding; TAS.
 DR GO: GO:0008334; P:histone mRNA metabolism; TAS.
 DR GO: GO:0006400; P:RNA modification; TAS.
 DR InterPro: IPR002344; Lupus_La.
 DR InterPro: IPR006630; Lupus_La_dom.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF05383; La; 1.
 DR Pfam: PF00076; trm; 1.
 DR PRINTS: PR00302; LUPUSLA.
 DR SMART: SM00715; LA; 1.
 DR SMART: SM00360; RRM; 1.
 DR PROSITE: PS50102; RRM; 1.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:01:41 ; Search time 7.224022 Seconds
(without alignments)
129.452 Million cell updates/sec

Title: US-09-836-073-4

Perfect score: 102
Sequence: 1 QOQEKICHQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	85.3	404	1 LA_BOVIN	P10881 Bos taurus
2	87	85.3	408	1 LA_HUMAN	P05455 homo sapien
3	87	85.3	415	1 LA_MOUSE	P32067 mus musculu
4	87	85.3	415	1 LA_RAT	P38656 rattus norv
5	66	64.7	427	1 LAB_XENLA	P28049 xenopus lae
6	66	64.7	428	1 LAB_XENLA	P28048 xenopus lae
7	57	55.9	390	1 LA_DROME	P40796 drosophila
8	51	50.0	383	1 LA_AEDAL	Q6457 aedes albop
9	46	45.1	298	1 LAH1_SCHPO	P87058 schizosacch
10	45	44.1	505	1 TCMD_PENCR	Q43033 petroselinu
11	43.5	42.6	1037	1 ACRD_ECOLI	P24177 escherichia
12	43	42.2	926	1 MEI9_DROME	Q24087 drosophila
13	42	41.2	242	1 GINO_BACST	P27675 bacillus st
14	41	40.2	251	1 FOL2_MOUSE	P05685 mus musculu
15	41	40.2	539	1 Z173_HUMAN	Q12899 homo sapien
16	41	40.2	539	1 Z173_PANTR	Q12899 pan troglod
17	41	40.2	1220	1 DPOL_HSVB	P28856 equine herp
18	41	40.2	1221	1 V143_NPVAC	P24307 autographa
19	40.5	39.7	1173	1 GIGA_ORYSA	Q64517 oryza sativ
20	40	39.2	155	1 RS15_HALMA	P05767 halocaula
21	40	39.2	249	1 VGS1_BPT4	P17173 bacterioph
22	40	39.2	466	1 SRO9_YEAST	P25567 saccharomyc
23	40	39.2	573	1 CATY_YEAST	P66115 saccharomyc
24	40	39.2	585	1 YH09_YEAST	P38818 saccharomyc
25	40	39.2	1107	1 MY1B_MOUSE	P46733 mus musculu
26	40	39.2	1136	1 MY1B_RAT	Q05096 rattus norv
27	39	38.2	403	1 KIM2_SHEEP	P25699 ovis aries
28	39	38.2	404	1 KIM1_HUMAN	O76009 homo sapien
29	39	38.2	412	1 KIM1_SHEEP	P02534 ovib aries
30	39	38.2	478	1 NRFA_ECOLI	P32050 escherichia
31	39	38.2	478	1 NRFA_SALTI	Q8149 salmonella
32	39	38.2	478	1 NRFA_SALTY	Q8149 salmonella
33	39	38.2	479	1 PRL2_ARATH	Q39190 arabidopsis

34	39	38.2	482	1 IPT5_HUMAN	Q13325 homo sapien
35	39	38.2	488	1 RBL_OLITU	P14959 olithodisc
36	39	38.2	507	1 NRFA_WOLSTU	Q8145 wolfinella s
37	39	38.2	755	1 SEC6_RAT	Q62825 rattus norv
38	39	38.2	756	1 SEC6_HUMAN	Q60645 homo sapien
39	39	38.2	1209	1 THR_DROME	P42286 drosophila
40	39	37.7	1070	1 P1B_RAT	Q92110 rattus norv
41	38.5	37.3	200	1 NACP_ECOLI	P33932 escherichia
42	38	37.3	281	1 Y818_PYRAE	Q82951 pyrobaculum
43	38	37.3	394	1 KIM4_HUMAN	Q76011 homo sapien
44	38	37.3	404	1 KIM3_HUMAN	Q14525 homo sapien
45	38	37.3	416	1 KIM1_HUMAN	Q15323 homo sapien

ALIGNMENTS

RESULT 1
LA_BOVIN STANDARD; PRT; 404 AA.
ID LA_BOVIN
AC P10881;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
DE SSB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RX MEDLINE=89202037; PubMed=2468131;
RA Chan E.K.L., Sullivan K.F., Tan E.M.;
RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences for RNA-binding."
RT Nucleic Acids Res. 17:2233-2244(1989).
RL
CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' terminl of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S, and 7-2 RNAs.
CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE C-TERMINAL PART OF THE PROTEIN.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; X13698; CAA31986.1; -.
CC PIR; S03849; S03849.
CC InterPro; IPR002344; Lupus La.
CC InterPro; IPR006630; Lupus_La_dom.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF05383; La; 1.
CC Pfam; PF00076; rrm; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00715; LA; 1.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS50102; RRM; 1.
CC PROSITE; PS00030; RRM_RNP_1; 1.
CC RNA-binding; Nuclear protein; Phosphorylation.

This Page Blank (uspto)

C;Species: *Bacillus halodurans*
 C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C;Accession: H83743
 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
 A;Reference number: A83650, PMID:20512582; PMID:11058132
 A;Accession: H83743
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-642 <STO>
 A;Cross-references: GB:AP001509; GB:BA000004; NID:G10173176; PIDN:BA04471.1; GSPDB:GN00
 A;Experimental source: strain C-125
 C;Genetic8:
 A;Gene: BH0752

Query Match 45.1%; Score 46; DB 2; Length 642;
 Best Local Similarity 56.2%; Pred. No. 14;
 Matches 9; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 2 QQEAKEICHOIEYFGD 17
 ::::|::|::|::|
 Db 471 QERKALVH--EYFEGD 484

RESULT 15

AE2408
 hypothetical protein alr4821 [imported] - Nostoc sp. (strain PCC 7120)
 C;Species: Nostoc sp. PCC 7120
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C;Accession: AE2408
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi
 Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
 DNA Res. 8, 205-213, 2001
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ana*
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Accession: AE2408
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-352 <KUR>
 A;Cross-references: GB:BA000019; PIDN:BA076520.1; PID:G17133958; GSPDB:GN00179
 A;Experimental source: strain PCC 7120
 C;Genetic8:
 A;Gene: alr4821

Query Match 44.1%; Score 45; DB 2; Length 352;
 Best Local Similarity 56.2%; Pred. No. 12;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 QQEAKEICHOIEYFG 16
 ::::|::|::|::|
 Db 172 QQERAKATOLIKYIG 187

Search completed: September 10, 2004, 18:02:28
 Job time : 15.0782 secs

RESULT 10

T00677 hypothetical protein At2g43970 [imported] - Arabidopsis thaliana

N:Alternate names: hypothetical protein F6B13.10

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 23-Mar-2001

C/Accession: T00677, G84872

R:Rounstley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,

submitted to the EMBL Data Library, June 1998

A:Description: Arabidopsis thaliana chromosome II BAC F6B13 genomic sequence.

A:Reference number: Z14180

A/Accession: T00677

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-529 <ROU>

A/Cross-references: EMBL:AC004005; NID:G3212846; PID:G3212854

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounstley, S.D.; Shea, T.P.; Bentlo, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Venkkes, S.E.; Umayam, L.; Tallon, L.

elms, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.

Nature 402, 761-768, 1999

A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: G84872

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-529 <STO>

A/Cross-references: GB:AE002093; NID:G3212854; PIDN:AAC23405.1; GSPDB:GN00139

C:Genetics:

A:Gene: F6B13.10; At2g43970

A/Map position: 2

A:introns: 200/3; 228/1; 259/3; 283/3; 310/2; 344/3; 367/3; 421/3

Query Match

Best Local Similarity 47.1%; Score 48; DB 2; Length 529;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 196 KIVNQVEYFSD 207

RESULT 11

T30953 hypothetical protein C44B4.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000

C/Accession: T30953

R:Sammons, L.; Wohldmann, P.; Gilliam, B.

submitted to the EMBL Data Library, August 1999

A:Description: The sequence of C. elegans cosmid C44B4.

A:Reference number: Z20945

A/Accession: T30953

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-396 <SAM>

A/Cross-references: EMBL:AF001140; PIDN:AAB54169.1

A:Experimental source: strain Bristol N2; clone C44B4

C:Genetics:

A:Map position: 1

A:introns: 45/1; 114/3

A/Note: C44B4.4

C:Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology

Query Match

Best Local Similarity 46.1%; Score 47; DB 2; Length 396;

Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 13 DQKIKQVEYFSD 26

RESULT 12

T38937 rna binding protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C/Accession: T38937

R:Badcock, K.; Churche, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, April 1997

A:Reference number: Z21818

A/Accession: T38937

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-298 <BAD>

A/Cross-references: EMBL:Z94864; PIDN:CAB08173.1; GSPDB:GN00066; SPDB:SPAC57A10.10C

A:Experimental source: strain 972h-; cosmid C57A10

C:Genetics:

A:Gene: SPDB:SPAC57A10.10C

A/Map position: 1

A:introns: 72/1

Query Match

Best Local Similarity 45.1%; Score 46; DB 2; Length 298;

Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 64 EAEVLKQVEYFSD 77

RESULT 13

T43542 RNA-binding protein la1 homolog - fission yeast (Schizosaccharomyces pombe)

N:Alternate names: la autocatalytic; ribonucleoprotein la homolog

C:Species: Schizosaccharomyces pombe

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000

C/Accession: T43542; T43325

R:Van Horn, D.J.; Yeo, C.J.; Xue, D.; Shi, H.; Wolin, S.L.

RNA 3, 1434-1443, 1997

A>Title: The la protein in Schizosaccharomyces pombe: a conserved yet dispensable phosph

A:Reference number: Z22560; MUID:98067398; PMID:9404894

A/Accession: T43542

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-298 <VAN>

A/Cross-references: EMBL:AF022949; PIDN:AAB82145.1

R:Utsumi, R.

submitted to the EMBL Data Library, February 1998

A:Description: Screening of S. pombe cDNA library using E. coli defective in signal tran

A:Reference number: Z22428

A/Accession: T43325

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-298 <UTS>

A/Cross-references: EMBL:AB011371; PIDN:BAA24981.1

C:Genetics:

A:Gene: slat

C:Function:

A:Description: the binding of the la protein to tRNA precursors is required for the endo

C:Keywords: phosphoprotein; RNA binding

Query Match

Best Local Similarity 45.1%; Score 46; DB 2; Length 298;

Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 64 EAEVLKQVEYFSD 77

RESULT 14

H83743 ABC transporter (permease) BR0752 [imported] - Bacillus halodurans (strain C-125)

R.Scherly, D.; Stutz, F.; Lin-Marq, N.; Clarkson, S.G.
J. Mol. Biol. 231, 196-204, 1993
A:Title: La proteins from *Xenopus laevis*: cDNA cloning and developmental expression.
A:Reference number: S33817; MUID:93287095; PMID:8510143
A:Accession: S33818
A:Molecule type: mRNA
A:Residues: 1-428 <CCH>
A:Cross-references: EMBL:X68817; NID:G64873; PIDN:CAA48715.1; PID:G64874
C:Comment: This protein associates with a variety of small RNA molecules, most of which
act as a transcription termination factor.
C:Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology
C:Keywords: phosphoprotein; RNA binding
F:112-178/Domain: ribonucleoprotein repeat homology <RRM>
F:113-118/Region: RNA-binding RNP2 motif
F:151-158/Region: RNA-binding RNP1 motif
F:428-428/Domain: phosphorylated #status predicted <PHY>

Query Match 64.7%; Score 66; DB 1; Length 428;
Best Local Similarity 78.6%; Pred. No. 0.005;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 EAKICHOIEYFGD 17
: ||| ||||| |||||
DB 14 DTKICEQIEYFGD 27

RESULT 6
La/SS-B homolog D-1a - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
C:Accession: A53773
R:Bal, C.; Li, Z.; Tollas, P.P.
Mol. Cell. Biol. 14, 5123-5129, 1994
A:Title: Developmental characterization of a *Drosophila* RNA-binding protein homologous to
A:Reference number: A53773; MUID:94309632; PMID:8035794
A:Accession: A53773
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-390 <BAI>
A:Cross-references: GB:U07652; NID:G464019; PIDN:AAA20518.1; PID:G464020
C:Genetics:
A:Gene: FlyBase:La
A:Cross-references: FlyBase:FBgn0011638
C:Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology
C:Keywords: leucine zipper; RNA binding

Query Match 55.9%; Score 57; DB 2; Length 390;
Best Local Similarity 62.5%; Pred. No. 0.14;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 QOEAKICHOIEYFGD 17
: ||| ||||| |||||
DB 49 KQERAILRQVEYFGD 64

RESULT 7
A53781
ribonucleoprotein la - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*
C:Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 20-Sep-1999
C:Accession: A53781
R:Yoo, C.U.; Wolin, S.L.
Mol. Cell. Biol. 14, 5412-5424, 1994
A:Title: La proteins from *Drosophila melanogaster* and *Saccharomyces cerevisiae*: a yeast
A:Reference number: A53781; MUID:94309661; PMID:8035818
A:Accession: A53781
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-390 <YOO>
A:Cross-references: GB:U32968; NID:G488469; PID:G488470
C:Genetics:
A:Gene: FlyBase:La

A:Cross-references: FlyBase:FBgn0011638
C:Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology
C:Keywords: RNA binding

Query Match 55.9%; Score 57; DB 2; Length 390;
Best Local Similarity 62.5%; Pred. No. 0.14;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 QOEAKICHOIEYFGD 17
: ||| ||||| |||||
DB 49 KQERAILRQVEYFGD 64

RESULT 8
H89777
capsular polysaccharide synthesis enzyme CapSP [imported] - *Staphylococcus aureus* (stra
C:Species: *Staphylococcus aureus*
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: H89777
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogur
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: H89777
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-391 <KUR>
A:Cross-references: GB:BA000018; PID:G13700080; PIDN:BA841379.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: capP
C:Superfamily: lipopolysaccharide biosynthesis protein bplD

Query Match 51.0%; Score 52; DB 2; Length 391;
Best Local Similarity 60.0%; Pred. No. 0.91;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 QOEAKICHOIEYFG 16
: ||| ||||| |||||
DB 362 QASRRICEAIEYFG 376

RESULT 9
T32701
hypothetical protein C14C6.12 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T32701
R:David, M.; Wohldmann, P.; Bauer, C.; Antoniou, B.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid C14C6.
A:Reference number: Z21210
A:Accession: T32701
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-166 <DAV>
A:Cross-references: EMBL:AF039051; PIDN:AB94258.1; GSPDB:GN00023; CESP:C14C6.12
A:Experimental source: strain Bristol N2; clone C14C6
C:Genetics:
A:Gene: CESP:C14C6.12
A:Map position: 5
A:Introns: 42/1; 156/3

Query Match 49.0%; Score 50; DB 2; Length 166;
Best Local Similarity 58.8%; Pred. No. 0.83;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 2 QOEAKICHOIEYFGD 18
: ||| ||||| |||||
DB 77 QDVKKICNVIEYMTGDF 93

A:Reference number: S03848; MUID:89202037; PMID:2468131
 A:Accession: S03848
 A:Molecule type: mRNA
 A:Residues: 1-408 <CH2>
 A:Cross-references: EMBL:X13697; NID:g36414; PIDN:CAA31985.1; PID:g36415
 R:Chambers, J.C.; Keene, J.D.
 Proc. Natl. Acad. Sci. U.S.A. 82, 2115-2119, 1985
 A:Title: Isolation and analysis of cDNA clones expressing human lupus La antigen.
 A:Reference number: A22956; MUID:85166283; PMID:3856888
 A:Accession: A22956
 A:Molecule type: mRNA
 A:Residues: 45-97, 'LK' <CH3>
 A:Cross-references: GB:J04205
 A>Note: this sequence has been revised in reference A31888
 R:Nyman, U.; Ringertz, N.R.; Petersson, I.
 Immunol. Lett. 22, 65-72, 1989
 A:Title: Demonstration of an amino terminal La epitope recognized by human anti-La sera.
 A:Reference number: A61051; MUID:89379261; PMID:2476379
 A:Accession: A61051
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-19, 'E', '21-47' <NTM>
 R:Stungess, A.D.; Peterson, M.G.; McNeillage, L.J.; Whittingham, S.; Coppel, R.L.
 J. Immunol. 140, 3212-3218, 1988
 A:Title: Characteristics and epitope mapping of a cloned human autoantigen La.
 A:Reference number: S11013; MUID:88199081; PMID:2452201
 A:Accession: S11013
 A:Molecule type: mRNA
 A:Residues: 'E', '55-287', 'V', '289-408' <STU>
 A:Cross-references: EMBL:M20328; NID:g337456; PIDN:AAA36577.1; PID:g337457
 R:Kohsaka, H.; Yamamoto, K.; Fujii, H.; Miura, H.; Miyasaka, N.; Nishioke, K.; Miyamoto, J.
 Clin. Invest. 85, 1566-1574, 1990
 A:Title: Fine epitope mapping the human SS-B/La protein: Identification of a distinct autoantigen.
 A:Reference number: I55553; MUID:90237237; PMID:1692037
 A:Accession: I55553
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 81-107 <RES>
 A:Cross-references: GB:M35261; NID:g338491; PIDN:AAA3652.1; PID:g338495
 A:Accession: I70205
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 174-224 <RE2>
 A:Cross-references: GB:M35263; NID:g338492; PIDN:AAA3653.1; PID:g338496
 A:Accession: I70206
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 279-342 <RE3>
 A:Cross-references: GB:M35262; NID:g338493; PIDN:AAA3654.1; PID:g338497
 C:Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.
 C:Gene: GDB:SSB
 A:Gene: GDB:SSB
 A:Cross-references: GDB:125359; OMIM:109090
 A:Map position: 2
 A:introns: 22/3; 57/2; 115/3; 151/3; 185/2; 209/2; 223/3; 264/3; 380/2
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:112-176/Domain: ribonucleoprotein repeat homology <RRM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:228-408/Domain: phosphorylated #status experimental <PHY>

Query Match 85.3%; Score 87; DB 1; Length 408;
 Best Local Similarity 100.0%; Pred. No. 1.7e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHOIEYRGDF 18
 |||||
 |||||

DB 14 EAKICHOIEYRGDF 28

RESULT 3

JC1494
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 24-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: JC1494; S25145
 R:Semsei, I.; Troester, H.; Bartsch, H.; Schwemmler, M.; Igloi, G.L.; Bachmann, M.
 Gene 126, 265-268, 1993
 A:Title: Isolation of rat cDNA clones coding for the autoantigen SS-B/La: Detection of a
 A:Reference number: JC1494; MUID:93246255; PMID:7916708
 A:Accession: JC1494
 A:Molecule type: mRNA
 A:Residues: 1-415 <SEM>
 A:Cross-references: GB:X67859; NID:g55778; PIDN:CAA48043.1; PID:g55779
 A:Experimental source: liver
 C:Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:112-178/Domain: ribonucleoprotein repeat homology <RRM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:227-415/Domain: phosphorylated #status predicted <PHY>

Query Match 85.3%; Score 87; DB 1; Length 415;
 Best Local Similarity 100.0%; Pred. No. 1.7e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHOIEYRGDF 18
 |||||
 |||||

DB 14 EAKICHOIEYRGDF 28

RESULT 4
 S33817
 ribonucleoprotein La-B - African clawed frog
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33817; S28544
 R:Scherry, D.; Stutz, F.; Lin-Mary, N.; Clarkson, S.G.
 J. Mol. Biol. 231, 196-204, 1993
 A:Title: La proteins from Xenopus laevis. cDNA cloning and developmental expression.
 A:Reference number: S33817; MUID:93287095; PMID:8510143
 A:Accession: S33817
 A:Molecule type: mRNA
 A:Residues: 1-427 <SCH>
 A:Cross-references: EMBL:X68818; NID:g64875; PIDN:CAA48716.1; PID:g64876
 C:Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:111-177/Domain: ribonucleoprotein repeat homology <RRM>
 F:112-117/Region: RNA-binding RNP2 motif
 F:150-157/Region: RNA-binding RNP1 motif
 F:227-427/Domain: phosphorylated #status predicted <PHY>

Query Match 64.7%; Score 66; DB 1; Length 427;
 Best Local Similarity 78.6%; Pred. No. 0.005;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 EAKICHOIEYRGDF 17
 |||||
 |||||

DB 13 DKICHOIEYRGDF 26

RESULT 5
 S33818

ribonucleoprotein La-A - African clawed frog
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33818; S28545

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:33:11 ; Search time 14.0782 Seconds
(without alignments)
122.988 Million cell updates/sec

Title: US-09-836-073-4

Perfect score: 102
Sequence: 1 QOQEAKEHQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR 78: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	85.3	404	1 S03849	ribonucleoprotein
2	87	85.3	408	1 A31888	ribonucleoprotein
3	87	85.3	415	1 UC1494	ribonucleoprotein
4	66	64.7	427	1 S33817	ribonucleoprotein
5	66	64.7	428	1 S33818	ribonucleoprotein
6	57	55.9	390	2 A53773	La/SS-B homolog D-
7	57	55.9	390	2 A53781	ribonucleoprotein
8	52	51.0	391	2 H89777	capsular polysacch
9	50	49.0	166	2 T32701	hypothetical prote
10	48	47.1	529	2 T00677	hypothetical prote
11	47	46.1	396	2 T30953	hypothetical prote
12	46	45.1	298	2 T38937	rna binding protei
13	46	45.1	298	2 T35432	RNA-binding protei
14	46	45.1	642	2 H83743	ABC transporter (p
15	45	44.1	352	2 AE2408	hypothetical prote
16	45	44.1	506	2 T14907	trans-cinnamate 4-
17	45	44.1	568	2 C82379	response regulator
18	43.5	42.6	788	2 A71076	hypothetical prote
19	43.5	42.6	1037	2 B65022	acridiflavin resista
20	43.5	42.6	1037	2 D91045	amiflavinoside eff
21	43.5	42.6	1037	2 G85889	hypothetical prote
22	43	42.2	658	2 D96656	hypothetical prote
23	43	42.2	658	2 T19487	hypothetical prote
24	43	42.2	926	2 S58936	metotic recombinat
25	43	42.2	1131	2 T15617	hypothetical prote
26	42	41.2	242	2 A42478	glutamine transpor
27	42	41.2	401	2 AC2207	L-cysteine/cystine
28	42	41.2	416	2 S60034	keratin Hal' type
29	42	41.2	424	2 T46197	hypothetical prote

30	42	41.2	449	2 G84091	hypothetical prote
31	42	41.2	505	2 B90181	Na+/H+ antiporter
32	42	41.2	506	2 T41623	probable protein k
33	42	41.2	515	2 T43152	hypothetical prote
34	41.5	40.7	1428	2 T13926	probable protein p
35	41	40.2	51	2 G81188	hypothetical prote
36	41	40.2	172	2 F75491	conserved hypothet
37	41	40.2	172	2 F75491	erythrocyte membra
38	41	40.2	251	2 B40869	folate-binding pro
39	41	40.2	319	2 AF0228	probable transpos
40	41	40.2	469	2 C70357	hypothetical prote
41	41	40.2	506	2 F85016	probable RING zinc
42	41	40.2	704	2 AE2107	serine/threonine k
43	41	40.2	845	2 T34064	hypothetical prote
44	41	40.2	1220	1 DJBRC3	DNA-directed DNA p
45	41	40.2	1220	2 T42573	DNA-directed DNA p

ALIGNMENTS

RESULT 1
S03849
ribonucleoprotein Ia - bovine
N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
C:Species: Bos primigenius taurus (cattle)
C>Date: 31-Dec-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C:Accession: S03849
R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.
Nucleic Acids Res. 17, 2233-2244, 1989
A>Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences
A:Reference number: S03848; MUID:89202037; PMID:2468131
A:Accession: S03849
A:Molecule type: mRNA
A:Residues: 1-404 <CHA>
A:Cross-references: EMBL:X13698; NID:g755; PIDD:CAA31986.1; PID:g756
A>Note: part of this sequence was confirmed by protein sequencing
C:Comment: This protein associates with a variety of small RNA molecules, most of which
act as a transcription termination factor.
C:Superfamily: ribonucleoprotein Ia; ribonucleoprotein repeat homology
C:Keywords: blocked amino end; phosphoprotein; RNA binding
F:112-178/Domain: ribonucleoprotein repeat homology <RRM>
F:113-118/Region: RNA-binding RNP2 motif
F:151-158/Region: RNA-binding RNP1 motif
F:228-404/Domain: phosphorylated status predicted <PHY>
Query Match 85.3%; Score 87; DB 1; Length 404;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CQ 4 EAKICHQIEYFGDF 18
DB 14 EAKICHQIEYFGDF 28
RESULT 2
A31888
ribonucleoprotein Ia - human
N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B; Sjogren syndrome antigen
C:Species: Homo sapiens (man)
C>Date: 21-May-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C:Accession: A31888; S03848; A22956; A61051; S11013; I55553; I70205; I70206; A31273
R:Chambers, J.C.; Kenan, D.; Martin, B.U.; Keene, J.D.
J. Biol. Chem. 263, 18043-18051, 1988
A>Title: Genomic structure and amino acid sequence domains of the human Ia autoantigen.
A:Reference number: A31888; MUID:89053970; PMID:3192525
A:Accession: A31888
A:Molecule type: mRNA
A:Residues: 1-408 <CHA>
A:Cross-references: GB:004205; NID:g178666; PIDD:AAA51885.1; PID:g178667
R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.
Nucleic Acids Res. 17, 2233-2244, 1989
A>Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences

This Page Blank (uspto)


```

RESULT 13
ID Q9QUG9 PRELIMINARY; PRT; 608 AA.
AC Q9QUG9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Calcium-and diacylglycerol-regulated guanine nucleotide exchange
DE factor 1.
GN RASGRP2 OR CALDAG-GEFI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99007305; PubMed=9789079;
RA Kawasaki H., Springett G.M., Toki S., Canales J.J., Harlan P.,
RA Blumenstiel J.P., Chen E.J., Bany I.A., Mochizuki N., Ashbacher A.,
RA Matsuda M., Housman D.E., Graybiel A.M.;
RT "A Rap guanine nucleotide exchange factor enriched highly in the basal
RT ganglia."
RL Proc. Natl. Acad. Sci. U.S.A. 95:13278-13283(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Kawasaki H., Housman D.E., Graybiel A.M.;
RA Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Chen E.J., Bany I.A., Mochizuki N., Ashbacher A., Matsuda M.,
RA Housman D.E., Graybiel A.M.;
RT "A Novel Rap Guanine Nucleotide Exchange Factor Enriched in the Basal
RT Ganglia."
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP EMBL; U08117; A012742.1; -.
DR EMBL; AF081193; AAC79697.1; -.
DR HSSP; P28867; IPTO.
DR MGD; MGI:133849; Raegrp2.
DR GO; GO:0005059; F:calcium ion binding; IEA.
DR GO; GO:0005085; F:guanyl-nucleotide exchange factor activity; IEA.
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR000651; RaegRFN.
DR InterPro; IPR001895; RaegRF CDC25.
DR InterPro; IPR008937; Ras GEF.
DR Pfam; PF00130; DAG_PE-bind; 1.
DR Pfam; PF00036; etfand; 2.
DR Pfam; PF00617; RaegRF; 1.
DR SMART; SM00109; C1; 1.
DR SMART; SM00147; RaegRFN; 1.
DR SMART; SM00229; RaegRFN; 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR PROSITE; PS00018; EF_HAND; 2.
DR PROSITE; PS00019; RaegRF CAT; 1.
DR PROSITE; PS00212; RaegRF_NTER; 1.
KW Calcium; Calcium-binding
SQ SEQUENCE 608 AA; 69294 MW; DC717794CE12C2D1 CRC64;

```

```

Query Match 46.9%; Score 46; DB 11; Length 608;
Best Local Similarity 41.2%; Pred. No. 26;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 2 ALOAKICHOIQYFGQF 18
DB 70 SLQVKTCHLVRYWVSAF 86

```

```

RESULT 14
Q9KUL7

```

```

ID Q9KUL7 PRELIMINARY; PRT; 568 AA.
AC Q9KUL7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Response regulator.
GN VCA1086.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OC NCBI_TaxId=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.R., Peterson J.D., Unsay L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Baas S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Niernan W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Frazer C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC -!- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
DR EMBL; AE004434; AAF96979.1; -.
DR PIR; C82379; C82379.
DR HSSP; P52934; 10MP.
DR TIGR; VCA1086; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0000160; P:two-component signal transduction system [p. . .]; IEA.
DR InterPro; IPR001932; P2C-like.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF00072; response_reg; 1.
DR Prodom; PD000039; Response_reg; 1.
DR SMART; SM00331; P2C_SIG; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS00110; RESPONSE_REGULATORY; 1.
KW Phosphorylation; Sensory transduction; Complete proteome.
SQ SEQUENCE 568 AA; 64331 MW; 87A02A350D986F51 CRC64;

```

```

Query Match 45.9%; Score 45; DB 16; Length 568;
Best Local Similarity 42.9%; Pred. No. 35;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 3 LOAKICHOIQYFG 16
DB 156 LMEEWCHQVEHIFG 169

```

```

RESULT 15
ID Q9LJ02 PRELIMINARY; PRT; 928 AA.
AC Q9LJ02;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Similar to KIA0731 protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrharioideae; Oryzaceae; Oryza.
OC NCBI_TaxId=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC

```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Becker;
RX MEDLINE=9617891; PubMed=8606192;
RA Sau S., Lee C.Y.;
RT "Cloning of type 8 capsule genes and analysis of gene clusters for the
production of different capsular polysaccharides in Staphylococcus
aureus.";
RL J. Bacteriol. 178:2118-2126(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Becker;
RX MEDLINE=97197525;
RA Sau S., Sun J., Lee C.Y.;
RT "Molecular characterization and transcriptional analysis of type 8
capsule genes in Staphylococcus aureus.";
RL J. Bacteriol. 179:1614-1621(1997).
DR EMBL; U73374; BAB49445.1; -.
DR HSSP; P27828; 1F6D.
DR GO; GO:0008761; F:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
DR GO; GO:0009103; P:lipo-polysaccharide biosynthesis; IEA.
DR GO; GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
DR InterPro; IPR003331; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.
DR TIGRFAMs; TIGR00236; wecB; 1.
SQ SEQUENCE 391 AA; 44247 MW; 1E8D9FAA9BC76F0D CRC64;

Query Match 46.9%; Score 46; DB 2; Length 391;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIQYFG 16
Db 366 RICEAIEYFG 376

RESULT 11
P95709 PRELIMINARY; PRT; 391 AA.
AC P95709;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE CapsP.
GN CAPSP.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Reynolds and Newman;
RX MEDLINE=97388587; PubMed=9245821;
RA Sau S., Bhassin N., Wann E.R., Lee J.C., Foster T.J., Lee C.Y.;
RT "The Staphylococcus aureus allelic genetic loci for serotype 5 and 8
capsule expression contain the type-specific genes flanked by common
genes.";
RL Microbiology 143:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Reynolds and Newman;
RX MEDLINE=98101481; PubMed=9440531;
RA Kiser K.B., Lee J.C.;
RT "Staphylococcus aureus caps50 and capsP genes functionally complement
mutations affecting enterobacterial common-antigen biosynthesis in
Escherichia coli.";
RL J. Bacteriol. 180:403-406(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Reynolds and Newman;
RX MEDLINE=98125727; PubMed=9466251;
RA Bhassin N., Albus A., Michon F., Livolsi P.J., Park J.-S., Lee J.C.;
RT "Identification of a gene essential for O-acetylation of the
Staphylococcus aureus type 5 capsular polysaccharide.";
```

```
RL Mol. Microbiol. 27:9-21(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Reynolds and Newman;
RA Bagga N., Wann E.R., Foster T.J., Lee J.C.;
RT Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL; U81973; AAC46099.1; -.
DR HSSP; P27828; 1F6D.
DR GO; GO:0008761; F:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
DR GO; GO:0009103; P:lipo-polysaccharide biosynthesis; IEA.
DR GO; GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
DR InterPro; IPR003331; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.
DR TIGRFAMs; TIGR00236; wecB; 1.
SQ SEQUENCE 391 AA; 44328 MW; B0105F690DB7CF1D CRC64;

Query Match 46.9%; Score 46; DB 2; Length 391;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIQYFG 16
Db 366 RICEAIEYFG 376

RESULT 12
P95X57 PRELIMINARY; PRT; 391 AA.
AC P95X57;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Capsular polysaccharide synthesis enzyme CapsP.
GN CAPP OR SAV0164 OR SA0159.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECTIS=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,
RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hoshoyama A.,
RA Mizumori-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuuchi J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003358; BAB56326.1; -.
DR EMBL; AP003129; BAB41379.1; -.
DR PIR; H89777; H89777.
DR HSSP; P27828; 1F6D.
DR GO; GO:0008761; F:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
DR GO; GO:0009103; P:lipo-polysaccharide biosynthesis; IEA.
DR GO; GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
DR InterPro; IPR003331; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.
DR TIGRFAMs; TIGR00236; wecB; 1.
KW Complete proteome.
SQ SEQUENCE 391 AA; 44372 MW; DDDF5FA715BCCECC CRC64;

Query Match 46.9%; Score 46; DB 16; Length 391;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIQYFG 16
Db 366 RICEAIEYFG 376
```

```

AC 07X050;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DE OSJNB0046p18.10 protein.
DE OSJNB0046p18.10.
OC Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriactoidae; Oryzaceae; Oryza.
OX NCBI_TaxId=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Jia G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,
RA Jia Y., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y.,
RA Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W.,
RA Sheng H.H., Gu J.L., Chen S.T., Ni L., Zhu F.H., Han B., Feng Q.,
RA Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X., Liu Y.L., Mu J., Yu Z.,
RA Chen L., Fan D.L., Wang Q.J., Zhang L., Lu Y.Q., Yu S.L., Liu X.H.,
RA Lu T.T., Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,
RA Qian Y.M., Ying X., Zhou B., Chen Z.H., Hao P., Zhang U., Wu M.,
RA Zhang R.Q., Guan J.P., Hong G.F.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
SQ SEQUENCE 863 AA; 96838 MW; 8ACE2362CCF2F296 CRC64;

Query Match
Best Local Similarity 48.0%; Score 47; DB 10; Length 863;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 3 L0AKICHOIQYFYG 16
Db 380 LOSQVCHERRYFLG 393

RESULT 8
ID 080WC0 PRELIMINARY; PRT; 141 AA.
AC 080WC0;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DE 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Similar to Ras, guanyl releasing protein 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Mammary gland;
RC MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo W.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallaby S.J.,
RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Murzyn D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman W., Madan A., Rodighiero S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Touchman J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RA Strausberg R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DE EMBL; BC051474; AAH51474.1; -.
DR GO; GO:0005085; P:guanyl-nucleotide exchange factor activity; IEA.
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR InterPro; IPR000651; RASGEFN.
DR InterPro; IPR008937; Ras GER.
DR SMART; SM00229; RASGEFN_1.
DR PROSITE; PS00212; RASGER_NTER; 1.
SQ SEQUENCE 141 AA; 16186 MW; 242C314B00AD9191 CRC64;

Query Match
Best Local Similarity 46.9%; Score 46; DB 11; Length 141;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 2 AL0AKICHOIQYFQGF 18
Db 70 SL0MKTCHLVRYWISAF 86

RESULT 9
ID 080YN8 PRELIMINARY; PRT; 381 AA.
AC 080YN8;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Capsular polysaccharide synthesis enzyme Cap8P.
GN CAP8P OR MW0139.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxId=196620;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwano N., Asano K., Naito T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA."
RL Lancet 359:1819-1827(2002).
DR EMBL; AP004822; BAB94004.1; -.
DR GO; GO:0008761; P:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR GO; GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
DR InterPro; IPR003311; Epimerase_2.
DR Pfam; PF02350; Epimerase 2; 1.
DR TIGRFAMs; TIGR00236; wecB; 1.
KM Complete proteome.
SQ SEQUENCE 381 AA; 43106 MW; 125E4D5D1904779C CRC64;

Query Match
Best Local Similarity 46.9%; Score 46; DB 16; Length 381;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 6 K1CH0IQYFYG 16
Db 356 RICEA1EYFYG 366

RESULT 10
ID P72382 PRELIMINARY; PRT; 391 AA.
AC P72382;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Cap8P.
GN CAP8P.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxId=1280;

```

```

RESULT 5
ID Q7ZT10 PRELIMINARY; PRT; 401 AA.
AC Q7ZT10;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Similar to sjogren syndrome antigen B (Autoantigen Ia).
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC045392; AAF45392.1; -.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR InterPro: IPR002344; Lupus_La.
DR InterPro: IPR006630; Lupus_La_dom.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; Rrm_1.
DR PRINTS: PR00302; LUPUSLA.
DR SMART: SM00715; LA_1.
DR SMART: SM00360; RRM_1.
DR PROSITE: PS50102; RRM_1.
DR PROSITE: PS00030; RRM_RNP_1.
SQ SEQUENCE 401 AA; 46138 MW; 2EF022FDD3916291 CRC64;

Query Match 51.0%; Score 50; DB 13; Length 401;
Best Local Similarity 57.1%; Pred. No. 3.6;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LQAKICHQIOYRFG 16
Db 10 LKKVMEQIEYFYG 23

RESULT 6
ID Q9UL65 PRELIMINARY; PRT; 671 AA.
AC Q9UL65; 000538;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Guanine exchange factor MCG7 isoform 1 (F25B3.3 Kinase like
DE protein).
GN CALDAG-GEFI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RL MEDLINE=20493616; PubMed=10918068;
RA Clyde-Smith J., Silins G., Gartside M., Grimmond S., Etheridge M.,
RA Apolloni A., Hayward N., Hancock J.F.;
RT Characterization of RasGRP2, a Plasma Membrane-targeted, Dual
RT Specificity Ras/Rap Exchange Factor.;
RL J. Biol. Chem. 275:32260-32267(2000).
[2]
RP SEQUENCE OF 63-671 FROM N.A.
RX MEDLINE=98001089; PubMed=9341881;
RA Kedra D., Seroussi E., Fransson I., Trifunovic J., Clark M.,
RA Lagercrantz J., Blennow E., Mehlin H., Dumanowski J.;
RT The germinal center kinase gene and a novel CDC25-like gene are
RT located in the vicinity of the PYGM gene on 11q13.;
RL Hum. Genet. 100:611-619(1997).
[3]

```

```

RP SEQUENCE OF 63-671 FROM N.A.
RX MEDLINE=99007305; PubMed=9789079;
RA Kawasaki H., Springett G.M., Toki S., Canales J.J., Harlan P.,
RA Blumenstiel J.P., Chen E.J., Bany I.A., Mochizuki N., Ashbacher A.,
RA Matsuda M., Housman D.E., Graybiel A.M.;
RT "A Rap guanine nucleotide exchange factor enriched highly in the basal
RT ganglia.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:13278-13283(1998).
[4]
RP SEQUENCE OF 63-671 FROM N.A.
RA Kawasaki H., Housman D.E., Graybiel A.M.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE OF 63-671 FROM N.A.
RA Kawasaki H., Housman D.E., Graybiel A.M.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
[6]
RP SEQUENCE OF 63-671 FROM N.A.
RA Kawasaki H., Springett G.M., Toki S., Canales J.J., Blumenstiel J.P.,
RA Chen E.J., Bany I.A., Mochizuki N., Ashbacher A., Matsuda M.,
RA Housman D.E., Graybiel A.M.;
RT "A Novel Rap Guanine Nucleotide Exchange Factor Enriched in the Basal
RT Ganglia.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE OF 63-671 FROM N.A.
RA Silins G.U., Grimmond S., Hayward N.;
RT Characterisation of a Novel Nucleotide Exchange Factor.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF043722; AAF07219.1; -.
DR EMBL: Y12336; CAA73005.1; -.
DR EMBL: U78170; AAD12741.1; -.
DR EMBL: AF081194; AAC79698.1; -.
DR EMBL: AF043723; AAF07220.1; -.
DR HSPSP; P28867; IPTQ.
DR Genew; HGNC:9879; RASGRP2.
DR GO: GO:0005509; F:calcium ion binding; IEA.
DR GO: GO:0019992; F:diacylglycerol binding; IEA.
DR GO: GO:0005085; F:guanyl-nucleotide exchange factor activity; IEA.
DR GO: GO:0008289; F:lipid binding; TAS.
DR GO: GO:0007165; P:signal transduction; TAS.
DR InterPro: IPR002219; DAG_PE-bind.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR000651; RasGEFN.
DR InterPro: IPR001895; RasGRF_CDC25.
DR Pfam: PF00130; DAG_PE-bind; 1.
DR Pfam: PF00036; ehand; 2.
DR Pfam: PF00617; RasGEF; 1.
DR SMART: SM00109; Cl_1.
DR SMART: SM00147; RasGEF; 1.
DR SMART: SM00229; RasGEFN; 1.
DR PROSITE: PS00479; DAG_PE_BIND_DOM_1; 1.
DR PROSITE: PS50081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE: PS00018; EF_HAND; 2.
DR PROSITE: PS50009; RASGEF_CAT; 1.
DR PROSITE: PS50212; RASGEF_NTER; 1.
DR Calcium; Calcium-binding; Phorbol-ester binding.
KV Calcium; Calcium-binding; Phorbol-ester binding.
SQ SEQUENCE 671 AA; 75547 MW; 67B7BD2B4FAEDAD CRC64;

Query Match 48.0%; Score 47; DB 4; Length 671;
Best Local Similarity 41.2%; Pred. No. 19;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ALQAKICHQIOYRFG 18
Db 132 SLQYKCHLVRYWISAF 148

RESULT 7
ID Q7XO50 PRELIMINARY; PRT; 863 AA.
[3]

```



```

DR GO:0005634; C.nucleus; IDA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR Pfam; PF00076; rtm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
SQ SEQUENCE 381 AA; 43891 MW; 2E2DEF1452C0F0E9 CRC64;

Query Match
Best Local Similarity 84.7%; Score 83; DB 11; Length 381;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALOAKICHOIQYFQGF 18
DB 12 ALEAKICHOIEYFGDF 28

RESULT 2
Q8BTU4 PRELIMINARY; PRT; 415 AA.
ID Q8BTU4;
AC Q8BTU4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sjogren syndrome antigen B.
GN SSB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002)
DR EMBL; AK088677; BAC40498.1; -.
DR MGD; MGI:98423; Ssb.
DR GO:0005634; C.nucleus; IDA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR Pfam; PF00076; rtm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 415 AA; 47657 MW; A7545C7686AC8363 CRC64;

Query Match
Best Local Similarity 84.7%; Score 83; DB 11; Length 415;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALOAKICHOIQYFQGF 18
DB 12 ALEAKICHOIEYFGDF 28

RESULT 3
Q8OHT5 PRELIMINARY; PRT; 206 AA.
ID Q8OHT5;
AC Q8OHT5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Untranslated region binding-protein.

```

```

GN UBP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA L'Ecuayer T.J., Fang H.-L.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF467897; AAL76269.1; -.
DR GO:0005634; C.nucleus; IEA.
DR GO:0003723; P:RNA binding; IEA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR Pfam; PF00076; rtm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 206 AA; 23992 MW; 965B62F7DFB90E9 CRC64;

Query Match
Best Local Similarity 63.3%; Score 62; DB 13; Length 206;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LOAKICHOIQYFQGF 16
DB 13 LDKICCOIEYFQGF 26

RESULT 4
Q7ZTK2 PRELIMINARY; PRT; 427 AA.
ID Q7ZTK2;
AC Q7ZTK2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to lupus LA protein homolog B.
DE Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC046654; AAH46654.1; -.
DR GO:0005634; C.nucleus; IEA.
DR GO:0003723; P:RNA binding; IEA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rtm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 427 AA; 48996 MW; 1E7CD82D8AB9C69A CRC64;

Query Match
Best Local Similarity 60.2%; Score 59; DB 13; Length 427;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LOAKICHOIQYFQGF 16
DB 12 LDKICCOIEYFQGF 25

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:29:01 ; Search time 41.229 Seconds
(without alignments)
137.751 Million cell updates/sec

Title: US-09-836-073-3

Perfect score: 98
Sequence: 1 AALQAKICHQIQYFQGF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	84.7	381	11	Q9CYB9 mus musculus
2	83	84.7	415	11	Q8BTU4 mus musculus
3	62	60.3	206	13	Q8OH15 gallus galli
4	59	60.2	427	13	Q7ZTK2 xenopus lae
5	50	51.0	401	13	Q7ZT10 brachydanio
6	47	48.0	671	4	Q9UL65 homo sapien
7	47	48.0	863	10	Q7XQ50 oryza sativ
8	46	46.9	141	11	Q8MWC0 mus musculus
9	46	46.9	381	16	Q8NVM8 staphylococ
10	46	46.9	391	2	P72382 staphylococ
11	46	46.9	391	2	P95708 staphylococ
12	46	46.9	391	16	Q99X57 staphylococ
13	46	46.9	608	11	Q9OU09 mus musculus
14	45	45.9	568	16	Q9K17 vibrio chol
15	45	45.9	928	10	Q9L102 oryza sativ
16	44	44.9	239	10	Q9SMH8 Q9smde laminiaria d

17	44	44.9	1676	8	Q85FR6	Q85f6 cyanidiosch
18	43	43.9	337	2	Q8RX98	Q8RX98 mycoplasma
19	42.5	43.4	577	16	Q97DB6	Q97db6 clostridium
20	42	42.9	179	5	Q814P8	Q814P8 plasmodium
21	42	42.9	188	5	Q81094	Q81094 plasmodium
22	42	42.9	244	5	Q81M79	Q81M79 plasmodium
23	42	42.9	396	5	Q01806	Q01806 caenorhabdi
24	42	42.9	478	16	Q87GS8	Q87gs8 vibrio para
25	42	42.9	541	16	Q9ZKY5	Q9Zky5 helicobacte
26	42	42.9	542	16	Q25534	Q25534 helicobacte
27	42	42.9	547	10	Q7XVC6	Q7Xvc6 oryza sativ
28	42	42.9	658	10	Q9CAN3	Q9can3 arabidopsis
29	42	42.9	1017	11	Q80Z25	Q80z25 mus musculu
30	42	42.9	1220	12	Q39272	Q39272 equine hepr
31	42	42.9	1249	12	Q65152	Q65152 african swi
32	42	42.9	1560	10	Q7XT99	Q7xt99 oryza sativ
33	42	42.9	2644	4	Q13535	Q13535 homo sapien
34	41	41.8	259	3	Q95K33	Q95k33 neurospora
35	41	41.8	271	16	Q8CTY0	Q8cty0 staphylococ
36	41	41.8	296	5	Q9W129	Q9w129 drosophila
37	41	41.8	336	2	Q50359	Q50359 mycoplasma
38	41	41.8	336	16	Q98QD2	Q98qd2 mycoplasma
39	41	41.8	366	2	Q85787	Q85787 mycoplasma
40	41	41.8	369	2	Q30382	Q30382 mycoplasma
41	41	41.8	369	2	Q9R8C2	Q9r8c2 mycoplasma
42	41	41.8	369	16	Q98EP1	Q98ep1 mycoplasma
43	41	41.8	399	2	Q85799	Q85799 mycoplasma
44	41	41.8	411	10	Q9FL36	Q9fl36 arabidopsis
45	41	41.8	422	10	Q94A38	Q94a38 arabidopsis

ALIGNMENTS

RESULT 1
Q9CYB9 PRELIMINARY; PRT; 381 AA.
AC Q9CYB9;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Sjogren eyndrome antigen B.
GN SSB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Atakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pease G., Quackenbush J.,
RA Schiraldi L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guernicich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima U., Mazzaletti U., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilting L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.",
RL Nature 409:685-690(2001).
DR EMBL; AK017822; BAB30957.1; -
MGD; MGI:98423; Seb.

This Page Blank (uspto)

FT METAL 18 18 IRON (HEME AXIAL LIGAND).
FT METAL 80 80 IRON (HEME AXIAL LIGAND).
SQ SEQUENCE 104 AA; 11209 MW; 1B5F2A7B657CD36F CRC64;

Query Match 39.8%; Score 39; DB 1; Length 104;
Best Local Similarity 75.0%; Pred. No. 8.3;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 QAKICHQI 11
DB 13 QCKICHQV 20

QY 2 ALQAKICHQIQ 12
DB 108 ALQOQCCHQIR 118

Search completed: September 10, 2004, 17:53:03
Job time : 8.24022 secs

RESULT 15

ZEB2_MAIZE
ID ZEB2_MAIZE STANDARD; PRT; 183 AA.
AC P08031;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Zein-beta precursor (Zein 2) (16 kDa) (Zein ZC1).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=87248094; PubMed=3596247;
RA Prat S., Perez-Grau L., Puigdemenech P.;
RT "Multiple variability in the sequence of a family of maize endosperm
RT proteins";
RL Gene 52:41-49(1987).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. W64A; TISSUE=Endosperm;
RX MEDLINE=91057131; PubMed=2243787;
RA Reina M., Guillen P., Ponte I., Boronat A., Palau J.;
RT "DNA sequence of the gene encoding the Zci protein from Zea mays W64
RT A.";
RL Nucleic Acids Res. 18:6425-6425(1990).
CC -1- FUNCTION: Zeins are major seed storage proteins.
CC -1- SUBCELLULAR LOCATION: Endosperm protein bodies.
CC -1- SIMILARITY: TO GLUTENIN 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M16460; AAA33523.1; -;
DR EMBL; X53515; CAA37595.1; -;
DR PIR; B29017; B29017.
DR MaizeDB; 58053; -;
DR InterPro; IPR003612; AAI.
DR InterPro; IPR001954; G1a_glutenin.
DR InterPro; IPR000480; Glutelin.
DR Pfam; PF00234; tryp_alpha_amy1; 1.
DR PRINTS; PR00208; GLIADGLUTEN.
DR PRINTS; PR00211; GLUTELIN.
DR SMART; SM00499; AAI; 1.
DR Seed storage protein; Repeat; Multigene family; signal.
FT SIGNAL 1 19
FT CHAIN 20 183 ZEIN-BETA.
SQ SEQUENCE 183 AA; 19558 MW; 3965BBBC1151F45A CRC64;
Query Match 39.8%; Score 39; DB 1; Length 183;
Best Local Similarity 63.6%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X82826; CAA58033.1; -
 DR EMBL; AB022217; BAB02756.1; ALT SEQ.
 DR EMBL; AY054181; AAL06842.1; ALT_INIT.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 7.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR PRODOM; PD000018; WD40; 2.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS00678; WD_REPEATS_1; 2.
 DR PROSITE; PS00082; WD_REPEATS_2; 4.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 FT REPEAT 168 198 WD 1.
 FT REPEAT 210 240 WD 2.
 FT REPEAT 252 282 WD 3.
 FT REPEAT 293 323 WD 4.
 FT REPEAT 335 364 WD 5.
 FT REPEAT 377 406 WD 6.
 FT REPEAT 426 456 WD 7.
 SQ SEQUENCE 479 AA; 53568 MW; 301986A4AER80670 CRC64;

Query Match
 Best Local Similarity 40.8%; Score 40; DB 1; Length 479;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIQYFG 16
 DB 48 RLCHKIQVAFG 58

RESULT 13
 TCPT_VIBCH STANDARD; PRT; 503 AA.
 AC T29480; Q9KTR0;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 16-OCT-2001 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Toxin coregulated pilus biosynthesis protein T (TCP pilus
 DE biosynthesis protein tcpt).
 GN TCPT OR VC0835.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxId=666;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=classical Inaba Z17561 / Serotype O1;
 RA Heidelberg J.F., Bisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Baise S., Olin H., Dragoi I., Sellers P.,
 RA McDonald L., Uetebach T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Frazer C.M.;
 RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae.";

RL Nature 406:477-483 (2000).
 CC - FUNCTION: INVOLVED IN THE TRANSLOCATION OF THE TCPA PILIN.
 CC - SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC - SIMILARITY: BELONGS TO THE PILE/OUTE/EXEE/XPSE/XCPR FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X64098; CAA54642.1; -
 DR EMBL; AE004168; AAF93998.1; -
 DR PIR; F82275; F82275.
 DR TIGR; VC0835; -
 DR InterPro; IPR001482; GSP1I_E.
 DR Pfam; PF00437; GSP1I_E; 1.
 DR PRODOM; PD000739; GSP1I_E; 1.
 DR PROSITE; PS00662; T2SP_E; 1.
 KW Transport; ATP-binding; Complete proteome.
 NP BIND 236 243 ATP (POTENTIAL).
 FT CONFLICT 137 137 A -> S (IN REF. 1).
 SQ SEQUENCE 503 AA; 57276 MW; ED3FDF8F579F916 CRC64;

Query Match
 Best Local Similarity 40.3%; Score 39.5; DB 1; Length 503;
 Matches 7; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 3 LQAKICHOQ-IQYFQ 17
 DB 471 LKGVCHETMLHFG 486

RESULT 14
 CY2_RHOVA STANDARD; PRT; 104 AA.
 AC P00082;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cytochrome C2.
 OS Rhodocrobium vanielii.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhodimicrobiaceae; Rhodimicrobium.
 OX NCBI_TaxId=1069;
 OX [1]
 RP SEQUENCE.
 RC STRAIN=ATCC 17100;
 RA MEDLINE=76102814; PubMed=174109;
 RA Ambler R.P., Meyer T.E., Kamen M.D.;
 RT "Primary structure determination of two cytochromes C2: close
 RT similarity to functionally unrelated mitochondrial cytochrome C.";
 RL Proc. Natl. Acad. Sci. U.S.A. 73:472-475 (1976).
 CC - FUNCTION: CYTOCHROME C2 IS FOUND MAINLY IN PURPLE, NONSULFUR,
 CC PHOTOSYNTHETIC BACTERIA WHERE IT FUNCTIONS AS THE ELECTRON DONOR
 CC TO THE OXIDIZED BACTERIOCHLOROPHYLL IN THE PHOTOPHOSPHORYLATION
 CC PATHWAY. HOWEVER, IT MAY ALSO HAVE A ROLE IN THE RESPIRATORY CHAIN
 CC AND IS FOUND IN SOME NONPHOTOSYNTHETIC BACTERIA.
 CC - PTM: Binds 1 heme group per molecule.
 DR PIR; A00074; CCRD2.
 DR HSSP; P00083; ICRT.
 DR InterPro; IPR003088; Cyt_C1.
 DR InterPro; IPR002327; Cyt_C1AB.
 DR InterPro; IPR000345; CytC_heme_BS.
 DR Pfam; PF00034; Cytochrome C; 1.
 DR PRINTS; PR00604; CYTCHEMCTAB.
 DR PRODOM; PD000375; Cyt_C1AB; 1.
 DR PROSITE; PS00190; CYTOCHROME C; 1.
 KW Electron transport; Photosynthesis; Heme.
 FT BINDING 14 14 HEME (COVALENT).
 FT BINDING 17 17 HEME (COVALENT).

YE14 SCHPO STANDARD; PRT; 646 AA.
 ID YE14 SCHPO
 AC 013659;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative GTP-binding protein CIB3.04c.
 GN SPAC1B3.04c.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OC NCBI_TaxId=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RC MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Heltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voicakeart G., Aert R., Robben J., Grymptre B.,
 RA Welfens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Beyer P., Zimmermann W., Wedler H., Mambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hunt S.M.,
 RA Lucas M., Rocher M., Galliard C., Tallada V.A., Garzon A., Rhode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,
 RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shipkova G.V., Uesery D., Barell B.G., Nure P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: Belongs to the GTP-binding elongation factor family.
 CC Lepa subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z98598; CAB1233.1; -
 CC PIR: T38022.
 CC HSP: P13551; IFNM.
 CC GeneDB: Spombe; SPAC1B3.04c; -
 CC InterPro: IPR000795; EF_GTPbind.
 CC InterPro: IPR000640; ERG_C.
 CC InterPro: IPR009022; ERG_III_V.
 CC InterPro: IPR004161; EFTU_D2.
 CC InterPro: IPR006297; LepA.
 CC InterPro: IPR005225; Small_GTP.
 CC InterPro: IPR009000; Translat_factor.
 CC Pfam: PF00679; ERG_C.1.
 CC Pfam: PF00009; GTP_EFTU.1.
 CC Pfam: PF0144; GTP_EFTU_D2.1.
 CC PRINTS: PR00315; ELONGATNCT.
 CC TIGRFAMs: TIGR001393; LepA.1.
 CC TIGRFAMs: TIGR00231; small_GTP.1.
 CC PROSITE: PS00301; EFATOR_GTP.1.
 CC Hypothetical protein: GTP-binding.
 KM

FT NE_BIND 65 72 GTP (POTENTIAL).
 FT NE_BIND 130 134 GTP (POTENTIAL).
 FT NE_BIND 184 187 GTP (POTENTIAL).
 SQ SEQUENCE 646 AA; 72683 MW; F9FA9498D384503E CRC64;
 Query Match 41.8%; Score 41; DB 1; Length 646;
 Best Local Similarity 46.7%; Freq. No. 25;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 Qy 3 LQAKICHOIQYVFGQ 17
 Db 108 VKAQCTSMIYYHGO 122
 RESULT 12
 PRL2_ARATH STANDARD; PRT; 479 AA.
 ID PRL2_ARATH
 AC Q39150; Q94007; Q9LUR9;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Pp1/PP2A phosphatases pleiotropic regulator PRL2.
 GN PRL2 OR A1361650 OR MGL6.11.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxId=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RC MEDLINE=98438452; PubMed=9765207;
 RX Nemeth K., Salchert K., Putnocky P., Bhalerao R., Koncz-Kalan Z.,
 RA Stankovic-Stangeland B., Bako L., Mathur J., Okresz L., Stabel S.,
 RA Getzenberger P., Sitt M., Redei G.P., Schnell J., Koncz C.;
 RT "Pleiotropic control of glucose and hormone responses by PRL1, a
 RT nuclear WD protein, in Arabidopsis.";
 RL Genes Dev. 12:3059-3073(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RC MEDLINE=20277480; PubMed=10819329;
 RX Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;
 RA "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 RA features of the regions of 4,504,864 bp covered by sixty P1 and TAC
 RT clones.";
 RL DNA Res. 7:131-135(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RC MEDLINE=22954850; PubMed=14593172;
 RX Yamada K., Lim U., Dale J.M., Chen H., Shinn P., Palm C.J.,
 RA Soultwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
 RA Kalish-Newman G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Tortum M.J.,
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Anestri P.,
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gujral M., Hansen N.F.,
 RA Hayashizaki Y., Johnson-Hopson C., Huan V.W., Iida K., Karnes M.,
 RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamita A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Empirical analysis of transcriptional activity in the Arabidopsis
 RT genome.";
 RL Science 302:842-846(2003).
 CC -1- FUNCTION: Pleiotropic regulator of Pp1 and PP2A phosphatases.
 CC -1- SIMILARITY: Contains 7 WD repeats.
 CC -1- SIMILARITY: Belongs to the WD-repeat PRL1/PRL2 family.
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
 CC gene model prediction.
 CC -1- CAUTION: It is uncertain whether Met-1 or Met-3 is the initiator.
 CC -----

QY 2 ALOAKICHOIYFGOF 18
 DB 8 AVKLKIRPIDYFSGF 24

RESULT 9
 IFITS_HUMAN STANDARD; PRT; 482 AA.

AC Q13325;
 DT 30-MAY-2000 (rel. 39, Created)
 DT 30-MAY-2000 (rel. 39, Last sequence update)
 DT 15-MAR-2004 (rel. 43, Last annotation update)
 DE Interferon-induced protein with tetratricopeptide repeats 5 (IFIT-5)
 DE (Retinotic acid- and interferon-inducible 58 kDa protein).
 GN IFIT5 OR R158.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98114590; PubMed=9398535;
 RA Nifkura T., Hirata R., Weil S.C.;
 RT "A novel interferon-inducible gene expressed during myeloid
 RT differentiation";
 RL Blood Cells Mol. Dis. 23:337-349(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas, and Spleen;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschuld S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stanclevon M., Soares M.B., Bonaldi M.F., Casavani T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallary S.J.,
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Maria M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- INDUCTION: By interferons.
 CC -1- SIMILARITY: Belongs to the IFIT family.
 CC -1- SIMILARITY: Contains 8 TPR repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, U34605; AAB84934.1; -
 CC EMBL, BC025786; AAB25786.1; -
 CC PIR, G02058; G02058.
 CC Genew, HGNC:13328; IFIT5.
 CC InterPro, IPR008940; Ptenyl_trans.
 CC InterPro, IPR001440; TPR.
 CC Pfam, PF00515; TPR; 5.
 CC SMART, SM00028; TPR; 5.
 CC Repeat, TPR repeat; interferon induction.
 CC REPEAT 51 84 TPR 1.
 CC REPEAT 94 127 TPR 2.

FT REPEAT 138 173 TPR 3.
 FT REPEAT 181 214 TPR 4.
 FT REPEAT 249 282 TPR 5.
 FT REPEAT 338 371 TPR 6.
 FT REPEAT 376 410 TPR 7.
 FT REPEAT 435 468 TPR 8.
 SQ SEQUENCE 482 AA; 55846 MW; 8045BC100384BE05 CRC64;

QY 9 HOIYOYFGOF 18
 DB 376 HOIHVYGRF 385

Query Match
 Best Local Similarity 42.9%; Score 42; DB 1; Length 482;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

RESULT 10
 DPOL_HSVB STANDARD; PRT; 1220 AA.

AC P28858;
 DT 01-DEC-1992 (rel. 24, Created)
 DT 01-DEC-1992 (rel. 24, Last sequence update)
 DT 28-FEB-2003 (rel. 41, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7).
 GN 30.
 OS Equine herpesvirus type 1 (strain Abap) (EHV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirinae.
 NC NCB1_TaxID=31520;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92295566; PubMed=1318606;
 RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
 RT "The DNA sequence of equine herpesvirus-1";
 RL Virology 189:304-316(1992).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + (DNA)(N).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Belongs to the DNA polymerase type-B family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, M86664; AAB02465.1; -
 CC PIR, D36798; D3BEC3.
 CC InterPro, IPR006172; DNA_pol_B.
 CC InterPro, IPR006134; DNA_pol_B_dom.
 CC InterPro, IPR006133; DNA_pol_B_exo.
 CC Pfam, PF00136; DNA_pol_B_1.
 CC Pfam, PF03104; DNA_pol_B_exo; 1.
 CC PRINTS, PR0106; DNAPOLB.
 CC SMART, SM00486; POLB; 1.
 CC PROSITE, PS00116; DNA_POLYMERASE_B; 1.
 CC Transferrase; DNA-directed DNA polymerase; DNA replication;
 CC DNA-binding; Nuclear protein.
 CC SEQUENCE 1220 AA; 135956 MW; 858C14DCCCT1A65B CRC64;

Query Match
 Best Local Similarity 42.9%; Score 42; DB 1; Length 1220;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 ICHOIYFG 16
 DB 248 VCHTLTYFG 257

RESULT 11

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; X68817; CAA48715.1; -.

DR PIR; S33818; S33818.

DR InterPro; IPR002344; Lupus_La.

DR InterPro; IPR006630; Lupus_La_dom.

DR InterPro; IPR000504; RNA_rec_mot.

DR Pfam; PF05383; La; 1.

DR Pfam; PF00076; rrm; 1.

DR PRINTS; PR00302; LUPUSLA.

DR SMART; SM00715; LA; 1.

DR SMART; SM00360; RRM; 1.

DR PROSITE; PS00102; RRM; 1.

DR PROSITE; PS00030; RRM_RNP_1; 1.

DR RNA-binding; Nuclear Protein; Phosphorylation.

FW DOMAIN 111 203 RNA-BINDING (RRM).

FT DOMAIN 196 212 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

FT DOMAIN 316 332 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

SQ SEQUENCE 428 AA; 48864 MW; AEB3A38B7D2B3EC3 CRC64;

Query Match 56.1%; Score 55; DB 1; Length 428;
Best Local Similarity 81.8%; Pred. No. 0.061;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 KICHOIQYRFG 16
DB 16 KICEQIEYRFG 26

RESULT 7

LA_AEDAL STANDARD; PRT; 383 AA.

ID LA_AEDAL

AC Q26457;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).

OS Aedes albopictus (Forest day mosquito).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Aedes.

OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.

OX NCBI_TaxID=1160;

RN [1]

RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.

RX MEDLINE=96135233; PubMed=8551578;

RA Pardigon N., Straus J.H.;

RT "Mosquito homolog of the La autoantigen binds to Sindbis virus RNA.";

RT J. Virol. 70:1173-1181(1996).

CC -1- FUNCTION: May be involved in transcription termination by RNA polymerase III. Binds RNA and DNA. Binds to the 3' end of the minus strand of Sindbis virus RNA. This may be significant for Sindbis virus RNA replication.

CC -1- SUBCELLULAR LOCATION: Nuclear. Primarily nuclear, but significant amounts are present in the cytoplasm.

CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.

CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; S80954; AAB35931.1; -.

DR InterPro; IPR002344; Lupus_La.

DR InterPro; IPR006630; Lupus_La_dom.

DR InterPro; IPR000504; RNA_rec_mot.

DR Pfam; PF05383; La; 1.

DR Pfam; PF00076; rrm; 1.

DR PRINTS; PR00302; LUPUSLA.

DR SMART; SM00715; LA; 1.

DR SMART; SM00360; RRM; 1.

DR PROSITE; PS00102; RRM; 1.

DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.

KW RNA-binding; Nuclear Protein; DNA-Binding.

FT DOMAIN 141 228 RNA-BINDING (RRM).

FT DOMAIN 141 228

SQ SEQUENCE 383 AA; 44430 MW; 4E5CC8F21C40F452 CRC64;

Query Match 43.9%; Score 43; DB 1; Length 383;
Best Local Similarity 50.0%; Pred. No. 6.5;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 LQAKICHOIQYRFG 16
DB 43 LEASTIRQLEYRFG 56

RESULT 8

EX5A_BUCAL STANDARD; PRT; 602 AA.

ID EX5A_BUCAL

AC P57530;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Exodeoxyribonuclease V alpha chain (EC 3.1.11.5).

GN RECD OR BU455.

OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum symbiotic bacterium).

OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; OC Enterobacteriaceae; Buchnera.

OX NCBI_TaxID=118099;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Tokyo 1998;

RX MEDLINE=20445173; PubMed=10993077;

RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;

RT "Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS.";

RT Nature 407:81-86(2000).

CC -1- FUNCTION: EXHIBITS SEVERAL CATALYTIC ACTIVITIES, INCLUDING ATP-DEPENDENT UNWINDING AND DNA-DEPENDENT ATPASE ACTIVITIES.

CC STRAND CLEAVAGE OCCURS 5' TO 3' DURING THE UNWINDING OF DUPLEX DNA AT CHI SEQUENCES, WHICH LOCALLY STIMULATE RECOMBINATION (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage (in the presence of ATP) in either 5'- to 3'- or 3'- to 5'-direction to yield 5'-phosphooligonucleotides.

CC -1- SUBUNIT: Consist of three subunits; recB, recC and recD (by similarity).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AP001119; BAB13153.1; -.

DR InterPro; IPR003593; AAA_ATPase.

DR InterPro; IPR006344; RecD.

DR SMART; SM00382; AAA; 1.

DR TIGRfams; TIGR01447; recD; 1.

DR HydroLase; Nuclease; Exonuclease; Endonuclease; Helicase; ATP-binding; KW DNA repair; Complete proteome.

FT NP BIND 171 178 ATP (POTENTIAL).

FT NP BIND 171 178

SQ SEQUENCE 602 AA; 69494 MW; 340FPAE3BBA456059 CRC64;

Query Match 43.9%; Score 43; DB 1; Length 602;
Best Local Similarity 47.1%; Pred. No. 10;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL, L00993; AAA39415.1; -
 DR EMBL, BC003820; AAH03820.1; -
 DR EMBL, Y07951; CA69249.1; -
 DR MGD; MGI:98423; Ssb.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; xrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 DR RNA-binding; Nuclear protein; RNA-BINDING (RRM).
 FT DOMAIN 111 187
 SQ SEQUENCE 415 AA; 4756 MW; 2D75197692FDC933 CRC64;

Query Match 84.7%; Score 83; DB 1; Length 415;
 Best Local Similarity 82.4%; Pred. No. 8.4e-07;
 Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALOAKICHQIQYFGOF 18
 Db 12 ALBAKICHOIEYFGDF 28

RESULT 5
 ID LAB_XENLA STANDARD; PRT; 427 AA.
 AC P28049;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein homolog B (La ribonucleoprotein B) (La autoantigen
 DE homolog B).
 GN LAB1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocyte;
 RX MEDLINE=93287095; PubMed=8510143;
 RA Scherly D., Stutz F., Lin-Marg N., Clarkson S.G.;
 RT "La proteins from Xenopus laevis. cDNA cloning and developmental
 RT expression.";
 RL J. Mol. Biol. 231:196-204 (1993).
 CC -!- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' termin of virtually all nascent
 CC polymerase III transcripts (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- DEVELOPMENTAL STAGE: Barely detectable in stage I/II oocytes,
 CC accumulate in stage III/IV oocytes, then exhibit a roughly
 CC constant steady state level in mature oocytes, eggs, and early
 CC embryos.
 CC -!- PTM: Phosphorylated (Probable).
 CC -!- MISCELLANEOUS: There are two forms of La, Laa and Lab, in xenopus.
 CC -!- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
 CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL, X68818; CAA48716.1; -
 DR PIR, S33817; S33817.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; xrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 DR RNA-binding; Nuclear protein; RNA-BINDING (RRM).
 FT DOMAIN 110 202
 SQ SEQUENCE 427 AA; 48995 MW; 45F3146F8934A355 CRC64;

Query Match 60.2%; Score 59; DB 1; Length 427;
 Best Local Similarity 71.4%; Pred. No. 0.012;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LQAKICHQIQYFG 16
 Db 12 LDTKICEQIEYFG 25

RESULT 6
 ID LAB_XENLA STANDARD; PRT; 428 AA.
 AC P28048;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein homolog A (La ribonucleoprotein A) (La autoantigen
 DE homolog A).
 GN LAB1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocyte;
 RX MEDLINE=93287095; PubMed=8510143;
 RA Scherly D., Stutz F., Lin-Marg N., Clarkson S.G.;
 RT "La proteins from Xenopus laevis. cDNA cloning and developmental
 RT expression.";
 RL J. Mol. Biol. 231:196-204 (1993).
 CC -!- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' termin of virtually all nascent
 CC polymerase III transcripts (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- DEVELOPMENTAL STAGE: Barely detectable in stage I/II oocytes,
 CC accumulate in stage III/IV oocytes, then exhibit a roughly
 CC constant steady state level in mature oocytes, eggs, and early
 CC embryos.
 CC -!- PTM: Phosphorylated (Probable).
 CC -!- MISCELLANEOUS: There are two forms of La, Laa and Lab, in xenopus.
 CC -!- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
 CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

DR PROSITE: PS00030; RRM RNP 1; 1.
 KW Systemic lupus erythematosus; RNA-binding; Phosphorylation;
 FT Nucleic protein.
 FT DOMAIN 111 187 RNA-BINDING (RRM).
 FT MOD RES 366 366 PHOSPHORYLATION (BY CK2).
 SQ SEQUENCE 408 AA; 46837 MW; EC153C15F9187FC4 CRC64;

Query Match 88.8%; Score 87; DB 1; Length 408;
 Best Local Similarity 83.3%; Pred. No. 1,7e-07;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
 |||:|||||:|||||
 DB 11 AALEAKICHQIEYFGDF 28

RESULT 3
 LA_RAT STANDARD; PRT; 415 AA.

AC P38656;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
 DE homolog).
 GN SSB OR SS-B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93246255; PubMed=7916708;
 RA Semsei I., Troester H., Bartsch H., Schwemle M., Igloi G.L.,
 RA Bachmann M.;
 RT "Isolation of rat cDNA clones coding for the autoantigen SS-B/La:
 RT detection of species-specific variations";
 RL Gene 1261265-266(1993).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' terminus of virtually all nascent
 CC polymerase III transcripts. It is associated with precursor forms
 CC of RNA polymerase III transcripts including crRNA and 4.5S, 5S, 7S,
 CC and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: Phosphorylated (By similarity).
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL: X67859; CAA48043.1; -.
 DR PIR: JCI494; JCI494.
 DR InterPro: IPR002344; Lupus_La.
 DR InterPro: IPR006630; Lupus_La_dom.
 DR InterPro: IPR005054; RNA_rec_mot.
 DR Pfam: PF05383; La; 1.
 DR Pfam: PF00076; rrm; 1.
 DR PRINTS: PRO0302; LUPUSLA.
 DR SMART: SM00715; La; 1.
 DR SMART: SM00360; RRM; 1.
 DR PROSITE: PS0102; RRM; 1.
 DR PROSITE: PS00030; RRM_RNP_1; 1.
 KW RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 111 187 RNA-BINDING (RRM).
 SQ SEQUENCE 415 AA; 47777 MW; 033FP9CC1E475F98 CRC64;

Query Match 88.8%; Score 87; DB 1; Length 415;
 Best Local Similarity 83.3%; Pred. No. 1,7e-07;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
 |||:|||||:|||||
 DB 11 AALEAKICHQIEYFGDF 28

RESULT 4
 LA_MOUSE STANDARD; PRT; 415 AA.

AC P32067;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
 DE homolog).
 GN SSB OR SS-B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93203630; PubMed=8454877;
 RA Topfer F., Gordon T., McCluskey J.;
 RT "Characterization of the mouse autoantigen La (SS-B). Identification
 RT of conserved RNA-binding motifs; a putative ATP binding site and
 RT reactivity of recombinant protein with poly(U) and human
 RT autoantibodies";
 RL J. Immunol. 150:3091-3100(1993).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary gland;
 RX MEDLINE=92388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klapper R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshuler S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
 RA Diatchenko L., Marziani K., Farmer A.F., Rubin G.M., Hong L.,
 RA Sapichenko M., Soares M.B., Bonaldo A.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Yoshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson S.D., Mulhally S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Guarniere P.H.,
 RA Richards S., Wotley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fehey J., Helton E., Kettelman M., Madan A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley A.C., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez R.W., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalak U., Smalins D.E.,
 RA Schenck A., Schein J.E., Jones S.U.W., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RL [3]
 RP SEQUENCE OF 1-11 FROM N.A.

RA Groelz D., Bachmann M.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' terminus of virtually all nascent
 CC polymerase III transcripts. It is associated with precursor forms
 CC of RNA polymerase III transcripts including crRNA and 4.5S, 5S, 7S,
 CC and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: Phosphorylated (By similarity).
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

FT DOMAIN 111 187 RNA-BINDING (RM).
 SQ. SEQUENCE 404 AA; 46534 MW; 4EE30B5C262AD6A1 CRC64;
 Query Match 88.8%; Score 87; DB 1; Length 404;
 Best Local Similarity 83.3%; Pred. No. 1.7e-07;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
 |||:|||||:|||||
 DB 11 AALEAKICHQIEYFGDF 28

RESULT 2
 LA HUMAN STANDARD; PRT; 408 AA.
 ID P05455;
 AC 01-NOV-1988 (Rel. 09, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein (Sjogren syndrome type B antigen) (SS-B) (La
 ribonucleoprotein) (La autoantigen).
 GN SSB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Eureleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89202037; PubMed=2468131;
 RA Chan E.K.L., Sullivan K.F., Tan E.M.;
 RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus
 Nucleic Acids Res. 17:2233-2244(1989).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89053970; PubMed=3192525;
 RA Chambers J.C., Kenan D., Martin B.J., Keene J.D.;
 RT "Genomic structure and amino acid sequence domains of the human La
 autoantigen.";
 RT J. Biol. Chem. 263:18043-18051(1988).
 [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta, and Skeletal muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Diatchenko U., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Cassavani T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustun T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulyaly S.J.,
 RA Bosak S.A., McManus P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Murray D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Heiton B., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield V.S.N., Krzywicki M.I., Skalka U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Merra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [4]
 RP SEQUENCE OF 54-408 FROM N.A.
 RX MEDLINE=88199081; PubMed=2455201;
 RA Scutgess A.D., Peterson M.G., McNeillage L.J., Whittingham S.,
 RA Coppel R.S.;
 RT "Characteristics and epitope mapping of a cloned human autoantigen
 La.";
 RT J. Immunol. 140:3212-3216(1988).
 [5]

RP SEQUENCE OF 54-97 FROM N.A.
 RX MEDLINE=85166283; PubMed=3856888;
 RA Chambers J.C., Keene J.D.;
 RT "Isolation and analysis of cDNA clones expressing human lupus La
 antigen.";
 RT Proc. Natl. Acad. Sci. U.S.A. 82:2115-2119(1985).
 [6]
 RP FUNCTION.
 RX MEDLINE=89251617; PubMed=2470590;
 RA Gottlieb E., Steltz J.A.;
 RT "Function of the mammalian La protein: evidence for its action in
 RT transcription termination by RNA polymerase III.";
 RT EMBO J. 8:851-861(1989).
 [7]
 RP PHOSPHORYLATION.
 RX MEDLINE=97207017; PubMed=9054510;
 RA Fan H., Sakulich A.L., Goodier J.L., Zhang X., Qin J., Marate R.J.;
 RT "Phosphorylation of the human La antigen on serine 366 can regulate
 RT recycling of RNA polymerase III transcription complexes.";
 RT Cell 88:707-715(1997).
 [8]
 RP INTERACTION WITH DDX15.
 RX MEDLINE=22346609; PubMed=12458796;
 RA Fournaux M.A., Kolman M.J.M., Van der Heijden A., De Jong A.S.,
 RA Van Venrooij W.J., Pruyn G.J.M.;
 RT "The human La (SS-B) autoantigen interacts with DDX15/hPrp43, a
 RT putative DEAD-box RNA helicase.";
 RT RNA 8:1428-1443(2002).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' terminus of virtually all nascent
 CC polymerase III transcripts. It is associated with precursor forms
 CC of RNA polymerase III transcripts including crRNA and 4.5S, 5S, 7S,
 CC and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE
 CC C-TERMINAL PART OF THE PROTEIN.
 CC -1- PTM: The N-terminus is blocked.
 CC -1- DISEASE: Sera from patients with systemic lupus erythematosus
 CC often contain antibodies that react with the normal cellular
 CC La protein as if this antigen was foreign.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; X13697; CAA11985.1; -
 DR EMBL; J04205; AAA51885.1; -
 DR EMBL; BC001289; AAH01289.1; -
 DR EMBL; BC020818; AAH20818.1; -
 DR PIR; A31888; A31888.
 DR GeneW; HGNC:11316; SSB.
 DR MIM; 109090; -
 DR GO; GO:0030529; C:ribonucleoprotein complex; TAS.
 DR GO; GO:0003729; F:RNA binding; TAS.
 DR GO; GO:0000449; F:RNA binding; TAS.
 DR GO; GO:0008334; P:histone mRNA metabolism; TAS.
 DR GO; GO:0006400; P:RNA modification; TAS.
 DR InterPro; IPR002344; Lupus La.
 DR InterPro; IPR006630; Lupus La dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; xrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SMO0715; LA; 1.
 DR SMART; SMO0360; RRM; 1.
 DR PROSITE; PSS0102; RRM; 1.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:01:41 ; Search time 7.24022 Seconds
(without alignments)
129.452 Million cell updates/sec

Title: US-09-836-073-3

Perfect score: 98
Sequence: 1 AALQAKICHQIQYFGQF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	88.8	404	1 LA_BOVIN	P10881 Bos taurus
2	87	88.8	404	1 LA_HUMAN	P5455 homo sapien
3	87	88.8	415	1 LA_RAT	P38656 rattus norv
4	83	84.7	415	1 LA_MOUSE	P32067 mus musculu
5	59	60.2	427	1 LAB_XENLA	P28049 xenopus lae
6	55	56.1	428	1 LAA_XENLA	P28046 xenopus lae
7	43	43.9	383	1 LAA_AEDAL	Q26457 aedes albop
8	43	43.9	602	1 EXSA_BUCAL	P51530 buchera ap
9	42	42.9	482	1 IPT5_HUMAN	Q13325 homo sapien
10	42	42.9	1220	1 DPOL_HSEVB	P28858 equine hearp
11	41	41.8	646	1 VE14_SCHPO	O13869 echinosacch
12	40	40.8	479	1 PRL2_ARATH	Q39190 arabidopsis
13	39.5	40.3	503	1 TCPT_VIBCH	P29480 vibrio chol
14	39	39.8	104	1 CY2_RHOVA	P00082 rhodocicrob
15	39	39.8	183	1 ZEB2_MAIZE	P08031 zea mays (m
16	39	39.8	390	1 LA_DROME	P40796 drosophila
17	39	39.8	397	1 TRPB_NITBU	Q82w12 nitrosomona
18	39	39.8	631	1 VE1_HPV30	Q05112 human papil
19	39	39.8	735	1 VQ37_SCHPO	Q10432 schizosacch
20	39	39.8	926	1 ME19_DROME	Q24087 drosophila
21	39	39.8	1221	1 V143_NPVAC	P24307 autographa
22	38.5	39.3	940	1 ADA_DROME	P91926 drosophila
23	38.5	39.3	977	1 A2A1_HUMAN	O95782 h adapter-r
24	38.5	39.3	977	1 A2A1_MOUSE	P17426 m adapter-r
25	38	38.8	251	1 FOL2_MOUSE	Q05685 mus musculu
26	38	38.8	305	1 HEM6_VIBCH	O9kv74 vibrio chol
27	38	38.8	786	1 STS5_HUMAN	P51692 homo sapien
28	38	38.8	1088	1 STS5_HUMAN	O8te98 homo sapien
29	38	38.8	1107	1 MY1B_MOUSE	P46735 mus musculu
30	38	38.8	1136	1 MY1B_RAT	Q05096 rattus norv
31	38	38.8	1181	1 HAIR_RAT	P97609 rattus norv
32	38	38.8	1189	1 HAIR_HUMAN	O45593 homo sapien
33	38	38.8	1596	1 GLI3_MOUSE	Q61602 mus musculu

34	37	37.8	227	1 MOAR_KLEAE	P54794 klebsiella
35	37	37.8	231	1 MTNA_BACSU	O32028 bacillus su
36	37	37.8	243	1 SDGF_RAT	P24338 rattus norv
37	37	37.8	309	1 FDHE_PSEAE	O9hv00 pseudomonas
38	37	37.8	359	1 CKR3_MOUSE	P51678 mus musculu
39	37	37.8	397	1 CARA_BUCBP	P59576 buchera ap
40	37	37.8	433	1 THIC_FUSNN	O8r100 fusobacteri
41	37	37.8	463	1 Y863_SYNY3	F73754 synechocyst
42	37	37.8	464	1 CASB_KLEOX	Q48409 klebsiella
43	37	37.8	466	1 SRO9_YEAST	P25567 saccharomyc
44	37	37.8	480	1 GLGA_RHITR	O9euts rhizobium c
45	37	37.8	487	1 GATB_CHLVC	Q823w6 chlamydomoph

ALIGNMENTS

RESULT 1
LA_BOVIN STANDARD; PRT; 404 AA.
ID LA_BOVIN
AC P10881;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
DE homolog).
GN SSB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RX MEDLINE=89202037; PubMed=2468131;
RA Chan E.K.L., Sullivan K.F., Tan E.M.;
RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus
sequences for RNA-binding."
RL Nucleic Acids Res. 17:2233-2244(1989).
CC -1- FUNCTION: La protein plays a role in the transcription of RNA
polymerase III. It is most probably a transcription termination
factor. Binds to the 3' terminus of virtually all nascent
polymerase III transcripts. It is associated with precursor forms
of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S,
and 7-2 RNAs.
CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE
C-TERMINAL PART OF THE PROTEIN.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
or send an email to license@ebi.ac.uk).
CC EMBL: X13698; CAA31986.1; -
CC PIR: S03849; S03849.
CC InterPro: IPR002344; Lupus La.
CC InterPro: IPR006630; Lupus La.dom.
CC InterPro: IPR005054; RNA_rec_mot.
CC Pfam: PF05383; La; 1.
CC Pfam: PF00076; rrm; 1.
CC PRINTS: PR00302; LupusLa.
CC SMART: SM00715; La; 1.
CC SMART: SM00360; RRM; 1.
CC PROSITE: PS50102; RRM; 1.
CC PROSITE: PS00030; RRM_RNP_1; 1.
CC RNA-binding; Nuclear protein; Phosphorylation.

This Page Blank (uspto)

Matches 5; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 3 LQAKICHQIQYRGGF 18
| : : | : : | : |

Db 520 LETRVCNLSLKYSAMP 535

RESULT 15

DJBEC3

DNA-directed DNA polymerase (EC 2.7.7.7) - equine herpesvirus 1 (strain Ab4p)

C;Species: equine herpesvirus 1

A;Note: host Equus caballus (domestic horse)

C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 27-Oct-2003

C;Accession: D36798

R;RefSeq: E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.

submitted to GenBank, March 1992

A;Description: The DNA sequence of equine herpesvirus-1.

A;Reference number: A36805

A;Accession: D36798

A;Molecule type: DNA

A;Residues: 1-1220 <TEL>

A;Cross-references: GB:M86664; NID:G330791; PIDN:AAB02465.1; PID:G330822

R;RefSeq: E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.

Virology 189, 304-316, 1992

A;Title: The DNA sequence of equine herpesvirus-1.

A;Reference number: A41831; MUID:92295566; PMID:1318606

A;Contents: annotation; possible protein-coding frames

A;Note: neither amino acid nor nucleotide sequence is given

C;Genetics:

A;Gene: 30

C;Superfamily: DNA polymerase

C;Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 42.9%; Score 42; DB 1; Length 1220;

Best Local Similarity 60.0%; Pred. No. 63;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 ICHQIQYRGGF 16
| : : | : : | : |

Db 248 VCHTTLTYFG 257

Search completed: September 10, 2004, 18:02:27
Job time : 16.0782 secs

RESULT 10

T30953

hypothetical protein C44E4.4 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000

C:Accession: T30953

R:Sammons, L.; Wohldmann, P.; Gillam, B.

submitted to the EMBL Data Library, August 1999

A:Description: The sequence of *C. elegans* cosmid C44E4.

A:Reference number: Z20945

A:Accession: T30953

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-396 <SAM>

A:Cross-references: EMBL:AF003140; PIDN:AA54169.1

A:Experimental source: strain Bristol N2; clone C44E4

C:Genetics:

A:Map position: 1

A:introns: 45/1; 114/3

A>Note: C44E4.4

C:Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology

Query Match 42.9%; Score 42; DB 2; Length 396;

Best Local Similarity 63.6%; Pred. No. 21;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 6 KICHQIYVFG 16

DB 15 KIKQLEVFG 25

RESULT 11

G02058

retinoid acid- and interferon-inducible 58k protein R158 - human

C:Species: *Homo sapiens* (man)

C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 29-Sep-1999

C:Accession: G02058

R:Well, S.C.

submitted to the EMBL Data Library, August 1995

A:Reference number: H00758

A:Accession: G02058

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-482 <WEI>

A:Cross-references: EMBL:U34605; NID:g1144510; PIDN:AAA4934.1; PID:g1144511

C:Superfamily: interferon-induced 56k protein

Query Match 42.9%; Score 42; DB 2; Length 482;

Best Local Similarity 60.0%; Pred. No. 26;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 9 HOIYVFG 18

DB 376 HOIHYGRF 385

RESULT 12

H71887

hypothetical protein jhp0797 - *Helicobacter pylori* (strain J99)C:Species: *Helicobacter pylori*

A:Variety: strain J99

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 28-Jul-2000

C:Accession: H71887

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

; Ives, C.; Gibson, R.; Metberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: H71887

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-541 <ARN>

A:Cross-references: GB:AE001509; GB:AE001439; NID:g4155350; PIDN:AAD06373.1; PID:g415536

A:Experimental source: strain J99

C:Genetics:

A:Gene: jhp0797

C:Superfamily: *Helicobacter pylori* hypothetical protein jhp0797

Query Match 42.9%; Score 42; DB 2; Length 541;

Best Local Similarity 50.0%; Pred. No. 29;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 AALQKICHQIQ 12

DB 123 AEVDKICHQVE 134

RESULT 13

G64627

hypothetical protein HP0863 - *Helicobacter pylori* (strain 26695)C:Species: *Helicobacter pylori*

C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 28-Jul-2000

C:Accession: G64627

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.

; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodex, A.; McKenne

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weldman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, C

A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: G64627

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-542 <TON>

A:Cross-references: GB:AE000596; GB:AE000511; NID:g2313982; PIDN:AAD07917.1; PID:g231400

C:Superfamily: *Helicobacter pylori* hypothetical protein jhp0797

Query Match 42.9%; Score 42; DB 2; Length 542;

Best Local Similarity 50.0%; Pred. No. 29;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 AALQKICHQIQ 12

DB 123 AEVDKICHQVE 134

RESULT 14

D96656

hypothetical protein Fl6M19.21 [imported] - *Arabidopsis thaliana*C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: D96656

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luhrs, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Talton,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A:Reference number: A66141; MUID:21016719; PMID:11130712

A:Accession: D96656

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-658 <STO>

A:Cross-references: GB:AE005173; NID:g6598840; PIDN:AAF18695.1; GSPDB:GN00141

C:Genetics:

A:Gene: Fl6M19.21

A:Map position: 1

Query Match 42.9%; Score 42; DB 2; Length 658;

Best Local Similarity 31.2%; Pred. No. 35;

R/Scherly, D.; Sprutz, F.; Lin-Mary, N.; Clarkson, S.G.
 J. Mol. Biol. 231, 196-204, 1993
 A/Title: La proteins from *Xenopus laevis*. cDNA cloning and developmental expression.
 A/Reference number: S33817; MUID:93287095; PMID:8510143
 A/Accession: S33818
 A/Molecule type: mRNA
 A/Residues: 1-428 <SC>
 A/Cross-references: EMBL:X68817, NID:664873; PIDN:CAA48715.1; PID:664874
 C/Comment: This protein associates with a variety of small RNA molecules, most of which
 act as a transcription termination factor.
 C/Superfamily: ribonucleoprotein la; ribonucleoprotein repeat homology
 C/Keywords: phosphoprotein; RNA binding
 F/112-178/Domain: ribonucleoprotein repeat homology <RRM>
 F/113-118/Region: RNA-binding RNP2 motif
 F/151-158/Region: RNA-binding RNP1 motif
 F/228-428/Domain: phosphorylated status predicted <PHY>

Query Match 56.1%; Score 55; DB 1; Length 428;
 Best Local Similarity 81.8%; Pred. No. 0.13;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 KICHOIQYFG 16
 DB 16 KICEQIEYFG 26

RESULT 6
 H89777
 capsular polysaccharide synthesis enzyme Caps [imported] - *Staphylococcus aureus* (strain
 C/Species: *Staphylococcus aureus*
 C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C/Accession: H89777
 R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
 ma, A.; Mizutani, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A/Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
 A/Reference number: A89758; MUID:21311952; PMID:11418146
 A/Accession: H89777
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-391 <KUR>
 A/Cross-references: GB:BA000018; PID:G13700080; PIDN:BA041379.1; GSPDB:GN00149
 A/Experimental source: strain N315
 C/Genetics:
 A/Gene: capP
 C/Superfamily: lipopolysaccharide biosynthesis protein bplD

Query Match 46.9%; Score 46; DB 2; Length 391;
 Best Local Similarity 63.6%; Pred. No. 4.3;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIQYFG 16
 DB 366 RICAIETVFG 376

RESULT 7
 C82379
 response regulator VCA1086 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)
 C/Species: *Vibrio cholerae*
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 28-Apr-2003
 C/Accession: C82379
 R/Heldberg, J.F.; Bisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.D.; Dodson, R.J.;
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A/Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A/Reference number: A82035; MUID:20406833; PMID:10952301
 A/Accession: C82379
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-568 <HEI>

A/Cross-references: GB:AE004434; GB:AE003853; NID:99558531; PIDN:AAF69679.1; GSPDB:GN00
 A/Experimental source: serogroup O1; strain N16961; biotype El Tor
 C/Genetics:
 A/Gene: VCA1086
 A/Map position: 2
 C/Superfamily: response regulator, Hnr type; response regulator homology

Query Match 45.9%; Score 45; DB 2; Length 568;
 Best Local Similarity 42.9%; Pred. No. 9.1;
 Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 LQAKICHOIQYFG 16
 DB 156 LMEMCHQVHIFG 169

RESULT 8
 A84983
 exodeoxyribonuclease V (EC 3.1.11.5) 67 kD polypeptide [imported] - *Buchnera* sp. (strain
 C/Species: *Buchnera* sp.
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
 C/Accession: A84983
 R/Shigenobu, S.; Matanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
 Nature 407, 81-86, 2000
 A/Title: Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp.
 A/Reference number: A84930; MUID:20445173; PMID:10993077
 A/Accession: A84983
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-602 <STO>
 A/Cross-references: GB:AP000398; GSPDB:GN00144
 A/Experimental source: strain APS
 C/Genetics:
 A/Gene: recD; BU455
 C/Superfamily: exodeoxyribonuclease V 67k chain
 C/Keywords: hydrolase

Query Match 43.9%; Score 43; DB 2; Length 602;
 Best Local Similarity 47.1%; Pred. No. 21;
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 ALQAKICHOIQYFG 18
 DB 8 AVKLKIRPIDFYSG 24

RESULT 9
 D97337
 mismatch repair protein Muts-like ATPase [imported] - *Clostridium acetobutylicum*
 C/Species: *Clostridium acetobutylicum*
 C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
 C/Accession: D97337
 R/Nolling, U.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee
 J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C1
 A/Reference number: A96900; MUID:21359325; PMID:21359325
 A/Accession: D97337
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-577 <KUR>
 A/Cross-references: GB:AE001437; PIDN:AAK81487.1; PID:G15026659; GSPDB:GN00168
 A/Experimental source: *Clostridium acetobutylicum* ATCC824
 C/Genetics:
 A/Gene: CAC3563

Query Match 43.4%; Score 42.5; DB 2; Length 577;
 Best Local Similarity 38.1%; Pred. No. 25;
 Matches 8; Conservative 6; Mismatches 2; Indels 5; Gaps 1;

QY 3 LQAKICHOIQY-----YFG 18
 DB 53 MKAKIKHLNRYERFRKRWGEF 73

A:Reference number: S03848; UID:89202037; PMID:2468131
 A:Accession: S03848
 A:Molecule type: mRNA
 A:Residues: 1-108 <CH2>
 A:Cross-references: EMBL:X13697; NID:936414; PID:CAA31985.1; PID:936415
 R:Chambers, J.C.; Keene, J.D.
 Proc. Natl. Acad. Sci. U.S.A. 82, 2115-2119, 1985
 A:Title: Isolation and analysis of cDNA clones expressing human lupus Ia antigen.
 A:Reference number: A22956; UID:85166283; PMID:3856888
 A:Accession: A22956
 A:Molecule type: mRNA
 A:Residues: 45-97, 'LK' <CH3>
 A:Cross-references: GB:J04205
 A:Note: This sequence has been revised in reference A31888
 R:Nyman, U.; Ringertz, N.R.; Petersson, I.
 Immunol. Lett. 22, 65-72, 1989
 A:Title: Demonstration of an amino terminal Ia epitope recognized by human anti-Ia sera.
 A:Reference number: A61051; UID:89379261; PMID:2476379
 A:Accession: A61051
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-19, 'E', 21-47 <NYM>
 R:Sturges, A.D.; Peterson, M.G.; McNeilage, L.J.; Whittingham, S.; Coppel, R.L.
 J. Immunol. 140, 3212-3218, 1988
 A:Title: Characteristics and epitope mapping of a cloned human autoantigen Ia.
 A:Reference number: S11013; UID:88199081; PMID:2452201
 A:Accession: S11013
 A:Molecule type: mRNA
 A:Residues: 'E', 55-287, 'V', 289-408 <STU>
 A:Cross-references: EMBL:M0328; NID:9337456; PID:AAA36577.1; PID:9337457
 R:Kohanka, H.; Yamamoto, K.; Fujii, H.; Mura, H.; Miyasaka, N.; Nishio, K.; Miyamoto, J.
 Clin. Invest. 85, 1566-1574, 1990
 A:Title: Fine epitope mapping the human SS-B/La protein: Identification of a distinct autoantigen.
 A:Reference number: I55553; UID:90237237; PMID:1692037
 A:Accession: I55553
 A:Molecule type: mRNA
 A:Status: translated from GB/EMBL/DBJ
 A:Residues: 81-107 <RES>
 A:Cross-references: GB:M3561; NID:9338491; PID:AAA36652.1; PID:9338495
 A:Accession: I70205
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 174-224 <RE2>
 A:Cross-references: GB:M3563; NID:9338492; PID:AAA36653.1; PID:9338496
 A:Accession: I70206
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 279-342 <RE3>
 A:Cross-references: GB:M3562; NID:9338493; PID:AAA36654.1; PID:9338497
 A:Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.
 C:Genetics:
 A:Gene: GDB:SSB
 A:Cross-references: GDB:125359; OMIM:109090
 A:Map position: 2
 A:Introns: 22/3, 57/2, 115/3, 151/3, 185/2, 209/2, 223/3, 264/3, 380/2
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:112-178/Domain: ribonucleoprotein repeat homology <RRM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:228-408/Domain: phosphorylated #status experimental <PHY>

Query Match 88.8%; Score 87; DB 1; Length 408;
 Best Local Similarity 83.3%; Pred. No. 3.7e-07;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALAKICHOIYFGF 18
 |||:|||||:|||||
 Db 11 AALEAKICHOIYFGF 28

RESULT 3

JC1494
 ribonucleoprotein La - rat
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 24-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: JC1494; S25145
 R:Semel, I.; Troester, H.; Bartsch, H.; Schwemmler, M.; Igloi, G.L.; Bachmann, M.
 Gene 126, 265-268, 1993
 A:Title: Isolation of rat cDNA clones coding for the autoantigen SS-B/La: Detection of a
 A:Reference number: JC1494; UID:93246255; PMID:7916708
 A:Accession: JC1494
 A:Molecule type: mRNA
 A:Residues: 1-415 <SEM>
 A:Cross-references: GB:X67859; NID:955778; PID:CAA48043.1; PID:955779
 A:Experimental source: liver
 A:Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:112-178/Domain: ribonucleoprotein repeat homology <RRM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:227-415/Domain: phosphorylated #status predicted <PHY>

Query Match 88.8%; Score 87; DB 1; Length 415;
 Best Local Similarity 83.3%; Pred. No. 3.8e-07;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALAKICHOIYFGF 18
 |||:|||||:|||||
 Db 11 AALEAKICHOIYFGF 28

RESULT 4

S33817
 ribonucleoprotein La.B - African clawed frog
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33817; S28544
 R:Schery, D.; Stutz, F.; Lin-Marq, N.; Clarkson, S.G.
 J. Mol. Biol. 231, 196-204, 1993
 A:Title: La proteins from Xenopus laevis. cDNA cloning and developmental expression.
 A:Reference number: S33817; UID:93287095; PMID:8510143
 A:Accession: S33817
 A:Molecule type: mRNA
 A:Residues: 1-427 <SCH>
 A:Cross-references: EMBL:X68818; NID:964875; PID:CAA48716.1; PID:964876
 A:Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:111-177/Domain: ribonucleoprotein repeat homology <RRM>
 F:112-117/Region: RNA-binding RNP2 motif
 F:150-157/Region: RNA-binding RNP1 motif
 F:227-427/Domain: phosphorylated #status predicted <PHY>

Query Match 60.2%; Score 59; DB 1; Length 427;
 Best Local Similarity 71.4%; Pred. No. 0.027;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LQAKICHOIYFG 16
 |||:|||||:|||||
 Db 12 LQAKICHOIYFG 25

RESULT 5

S33818
 ribonucleoprotein La.A - African clawed frog
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33818; S28545

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:33:11 ; Search time 14.0782 Seconds
(without alignments)
122.988 Million cell updates/sec

Title: US-09-836-073-3

Perfect score: 98
Sequence: 1 AALQAKICHQIYYFGQF 18

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.78.*

1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	88.8	404	1	S03849
2	87	88.8	408	1	A31888
3	87	88.8	415	1	JC1494
4	59	60.2	427	1	S33817
5	55	56.1	428	1	S33818
6	46	46.9	391	2	H89777
7	45	45.9	568	2	C82379
8	43	43.9	602	2	A84983
9	42.5	43.4	577	2	D97337
10	42	42.9	396	2	T30953
11	42	42.9	482	2	G02058
12	42	42.9	541	2	H71887
13	42	42.9	542	2	G64627
14	42	42.9	658	2	D96656
15	42	42.9	1220	1	DJBEC3
16	42	42.9	1220	2	T42573
17	42	42.9	259	2	T49596
18	41	41.8	336	2	S49396
19	41	41.8	336	2	B90566
20	41	41.8	369	2	G90596
21	41	41.8	646	2	T38022
22	40	40.8	120	2	AT1482
23	40	40.8	133	2	A83987
24	40	40.8	175	2	S23697
25	40	40.8	380	2	UC5217
26	40	40.8	598	2	T05888
27	40	40.8	658	2	T19487
28	39.5	40.3	163	2	H87733
29	39.5	40.3	503	2	F82275

30	39.5	40.3	503	2	JN0524	lept protein - Vib
31	39	39.8	104	1	CCRD2	cytochrome c2 - Rh
32	39	39.8	166	2	T32701	hypothetical prote
33	39	39.8	183	2	B29017	zein Zc1 - maize
34	39	39.8	346	2	A82659	two-component syst
35	39	39.8	388	2	AF0183	probable exported
36	39	39.8	390	2	A53773	La/SS-B homolog D-
37	39	39.8	390	2	A53781	ribonucleoprotein
38	39	39.8	461	2	T27651	hypothetical prote
39	39	39.8	487	2	T27653	hypothetical prote
40	39	39.8	505	2	B90181	Na+/H+ antiporter
41	39	39.8	506	2	F85016	probable RING zinc
42	39	39.8	529	2	T00677	hypothetical prote
43	39	39.8	631	2	S36505	hypothetical prote
44	39	39.8	696	2	F83886	penicillin-binding
45	39	39.8	735	2	T41187	hypothetical prote

ALIGNMENTS

RESULT 1

S03849
ribonucleoprotein La - bovine
N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
C:Species: Bos primigenius taurus (cattle)
C>Date: 31-Dec-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C:Accession: S03849
R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.
Nucleic Acids Res. 17, 2233-2244, 1989
A>Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences
A:Reference number: S03848; MUID:89202037; PMID:2468131
A:Accession: S03849

A:Molecule type: mRNA
A:Residues: 1-404 <CH>
A:Cross-references: EMBL:X13698; NID:q755; PIDN:CAA31986.1; PID:q756
A:Note: part of this sequence was confirmed by protein sequencing
C:Comment: This protein associates with a variety of small RNA molecules, most of which
ay act as a transcription termination factor.
C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
C:Keywords: blocked amino end; phosphoprotein; RNA binding
F:112-178/Domain: ribonucleoprotein repeat homology <RHM>
F:113-118/Region: RNA-binding RNP2 motif
F:151-158/Region: RNA-binding RNP1 motif
F:228-404/Domain: phosphorylated #status predicted <PHX>

Query Match
Best Local Similarity 88.8%; Score 87; DB 1; Length 404;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Cy 1 AALQAKICHQIYYFGQF 18
Db 11 AALEAKICHQIYYFGDF 28

RESULT 2

A31888
ribonucleoprotein La - human
N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B; Sjogren syndrome antigen
C:Species: Homo sapiens (man)
C>Date: 21-May-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.
Nucleic Acids Res. 17, 2233-2244, 1989
A>Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences
A:Reference number: A31888; MUID:89053970; PMID:3192525
A:Accession: A31888
A:Molecule type: mRNA
A:Residues: 1-408 <CH>
A:Cross-references: GB:J04205; NID:q178686; PIDN:AAA51885.1; PID:q178687
R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.
Nucleic Acids Res. 17, 2233-2244, 1989
A>Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences

Db 278 LRAKILQVEYYFGSD 293
 Search completed: September 10, 2004, 17:59:59
 Job time : 44.229 secs

DT 01-JUN-2003 (T-EMBLrel. 24, last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, last annotation update)
 DE Putative GntR-family regulatory protein.
 GN VPA1237.
 OS Vibrio parahaemolyticus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrrio.
 OC NCBI_TaxID=670;
 OX NCBI_TaxID=670;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RIMD 2210633 / Serotype O3:K6;
 RX MEDLINE=22508454; Pubmed=12620739;
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
 RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
 RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.,
 RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
 RT distinct from that of V. cholerae.";
 RL Lancet 361:743-749(2003).
 DR EMBL: AP005088; BAC62580.1; -;
 DR GO: GO:0005622; C:intracellular; IEA.
 DR GO: GO:0003700; F:transcription factor activity; IEA.
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro: IPR000524; HTH_GntR.
 DR Pfam: PF00392; gntR; 1.
 DR PRINTS: PR00035; HTHGNT.
 DR PROSITE: PS00043; HTH_GNTR_FAMILY; 1.
 DR Complete proteome.
 KW SEQUENCE 478 AA; 53805 MW; 402B8754DE010DDE CRC64;
 SQ

Query Match 49.0%; Score 47; DB 16; Length 478;
 Best Local Similarity 58.3%; Pred. No. 21;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 6 QICQCIERYFGD 17
 Db 386 QVCFSIQHYFGD 397

RESULT 15
 Q9LJ02
 ID Q9LJ02 PRELIMINARY; PRT; 928 AA.
 AC Q9LJ02;
 DT 01-OCT-2000 (T-EMBLrel. 15, Created)
 DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE Similar to KIAA0731 protein.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriaraloideae; Oryzaceae; Oryza.
 OC NCBI_TaxID=4530;
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.,
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RT clone: P0499C11.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP001080; BAA90356.1; -;
 DR Gramene: Q9LJ02; -;
 DR InterPro: IPR006607; DUF_DM15.
 DR InterPro: IPR006630; Lupus_La_dom.
 DR Pfam: PF05383; La; 1.
 DR SMART: SM00684; DM15; 3.
 DR SMART: SM00715; LA; 1.
 SQ SEQUENCE 928 AA; 103745 MW; 53FC46E2A4446EB4 CRC64;

Query Match 47.4%; Score 45.5; DB 10; Length 928;
 Best Local Similarity 62.5%; Pred. No. 73;
 Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

OY 3 LEAQICQCIERYFGD 17

```

RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Onca T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-O., Ito T.,
RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hosoyma A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekintzu K., Hirakawa H., Kihara S., Goto S., Yabuzaki J.,
RA Kanetsu M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancel 357:1225-1240(2001).
DR EMBL, AP003358; BAB56326.1; -.
DR EMBL, AP003129; BAB41319.1; -.
DR FIR, H89777; H89777.
DR HSP, P27828; I6D.
DR GO: GO:0008761; P:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
DR GO: GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR GO: GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
DR InterPro: IPR003331; Epimerase_2.
DR Pfam: PF02350; Epimerase_2; 1.
DR TIGRFAMs: TIGR00236; wecB; 1.
KM Complete proteome.
SQ SEQUENCE 391 AA; 44372 MW; DDDF5FA715BCCECC CRC64;

Query Match 52.1%; Score 50; DB 16; Length 391;
Best Local Similarity 72.7%; Pred. No. 5.4;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 QICQIQIEYFPG 16
Db 366 RICEAIEYFPG 376

RESULT 12
Q9FLJ6 PRELIMINARY; PRT; 411 AA.
AC Q9FLJ6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similarity to RNA-binding protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=96344145; PubMed=9679202;
RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
RT features of the regions of 1,381,565 bp covered by twenty one
RT physically assigned pl and TAC clones."
RL DNA Res. 5:131-145(1998).
DR EMBL, AB010638; BAB11080.1; -.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR InterPro: IPR002344; Lupus_La.
DR InterPro: IPR006630; Lupus_La_dom.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF05383; La; 1.
DR Pfam: PF00076; rrm; 1.
DR PRINTS: PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
SQ SEQUENCE 411 AA; 45655 MW; A2EF62EB5B89B099 CRC64;

Query Match 49.0%; Score 47; DB 10; Length 411;
Best Local Similarity 53.3%; Pred. No. 18;

```

```

Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 3 LEAQICQIEYFPG 17
Db 92 LNKIRIQVEYFSD 106

RESULT 13
Q94A38 PRELIMINARY; PRT; 422 AA.
ID Q94A38;
AC Q94A38;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE A05946250/MPL12.3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN RP SEQUENCE FROM N.A.
RA Shin P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Bath J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamita A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shin P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
RA Hsuan V.W., Ishida J., Jones T., Kamita A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,
Ecker J.R.;
RT "Arabidopsis ORF clones."
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL, AY050403; AAY91419.1; -.
DR EMBL, BT000588; AAN18157.1; -.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR InterPro: IPR002344; Lupus_La.
DR InterPro: IPR006630; Lupus_La_dom.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF05383; La; 1.
DR Pfam: PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
SQ SEQUENCE 422 AA; 46842 MW; 4EC4BBF1E068F0E CRC64;

Query Match 49.0%; Score 47; DB 10; Length 422;
Best Local Similarity 53.3%; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 3 LEAQICQIEYFPG 17
Db 103 LNKIRIQVEYFSD 117

RESULT 14
Q87GS8 PRELIMINARY; PRT; 478 AA.
ID Q87GS8;
AC Q87GS8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

```

```

RT acquired MRSA";
RL Lancet 359:1819-1827(2002).
DR EMBL; AB004822; BAB94004.1;
DR GO; GO:0008761; F:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR GO; GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
DR InterPro; IPR003311; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.
DR TIGRFAMs; TIGR00236; wecB; 1.
KW Complete proteome.
SQ SEQUENCE 381 AA; 43106 MW; 125E4D5D1904779C CRC64;

Query Match
Best Local Similarity 52.1%; Score 50; DB 16; Length 381;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 6 OICQOIEYFG 16
DB 366 RICEAIEYFG 366

RESULT 9
ID P72382 PRELIMINARY; PRT; 391 AA.
AC P72382;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CAPS.
GN CAPS.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_Taxid=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Becker;
RX MEDLINE=96178981; PubMed=8606197;
RA Sau S., Lee C.Y.;
RT "Cloning of type 8 capsular genes and analysis of gene clusters for the
RT production of different capsular polysaccharides in Staphylococcus
RT aureus";
RT J. Bacteriol. 178:2118-2126(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Becker;
RX MEDLINE=97197525;
RA Sau S., Sun J., Lee C.Y.;
RT "Molecular characterization and transcriptional analysis of type 8
RT capsular genes in Staphylococcus aureus";
RT J. Bacteriol. 179:1614-1621(1997).
RL EMBL; U73374; AAB49445.1;
DR HSSP; P27828; 1F6D.
DR GO; GO:0008761; F:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR InterPro; IPR003311; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.
DR TIGRFAMs; TIGR00236; wecB; 1.
SQ SEQUENCE 391 AA; 44247 MW; 1E8D9FAA9BC76F0D CRC64;

Query Match
Best Local Similarity 52.1%; Score 50; DB 2; Length 391;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 6 OICQOIEYFG 16
DB 366 RICEAIEYFG 376

RESULT 10
ID P95709 PRELIMINARY; PRT; 391 AA.
AC P95709;

```

```

DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CAPS.
GN CAPS.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_Taxid=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Reynolds and Newman;
RX MEDLINE=97185887; PubMed=9245821;
RA Sau S., Bhasin N., Mann E.R., Lee J.C., Foster T.J., Lee C.Y.;
RT "The Staphylococcus aureus allelic genetic loci for serotype 5 and 8
RT capsular expression contain the type-specific genes flanked by common
RT genes";
RT J. Microbiol. 143:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Reynolds and Newman;
RX MEDLINE=98101481; PubMed=9440531;
RA Kiser K.B., Lee J.C.;
RT "Staphylococcus aureus caps5 and capsP genes functionally complement
RT mutations affecting enterobacterial common-antigen biosynthesis in
RT Escherichia coli";
RT J. Bacteriol. 180:403-406(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Reynolds and Newman;
RX MEDLINE=98125727; PubMed=9466251;
RA Bhasin N., Albus A., Michon F., Iyolai P.J., Park J.-S., Lee J.C.;
RT "Identification of a gene essential for O-acetylation of the
RT Staphylococcus aureus type 5 capsular polysaccharide";
RT Mol. Microbiol. 27:9-21(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Reynolds and Newman;
RA Baga N., Mann E.R., Foster T.J., Lee J.C.;
RT Submitted (DEC-1996) to the EMBL/Genbank/DBJ databases.
RL EMBL; U81973; AAC46099.1;
DR HSSP; P27828; 1F6D.
DR GO; GO:0008761; F:UDP-N-acetylglucosamine 2-epimerase activity; IEA.
DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR GO; GO:0006047; P:UDP-N-acetylglucosamine metabolism; IEA.
DR InterPro; IPR003311; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.
DR TIGRFAMs; TIGR00236; wecB; 1.
SQ SEQUENCE 391 AA; 44328 MW; B0105F690DB7CFID CRC64;

Query Match
Best Local Similarity 52.1%; Score 50; DB 2; Length 391;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 6 OICQOIEYFG 16
DB 366 RICEAIEYFG 376

RESULT 11
ID O99X57 PRELIMINARY; PRT; 391 AA.
AC O99X57;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Capsular polysaccharide synthesis enzyme CapsP.
GN CAPS OR SAV0164 OR SA0159.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_Taxid=158876; 158879;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RESULT 5
Q87T10 PRELIMINARY; PRT; 401 AA.
ID Q87T10
AC Q87T10;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to Sjogren syndrome antigen B (Autoantigen La).
OC Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_Taxid=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC045392; AAI45392.1; -.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR InterPro: IPR002344; Lupus_La.
DR InterPro: IPR006630; Lupus_La_dom.
DR InterPro: IPR00504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 1.
DR PRINTS: PR00302; LUPUSLA.
DR SMART: SM00715; LA; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PSS0102; RRM; 1.
DR PROSITE: PSS0030; RRM_NRP_1; 1.
SQ SEQUENCE 401 AA; 46138 MW; 2EF032FDD3916291 CRC64;

Query Match 62.5%; Score 60; DB 13; Length 401;
Best Local Similarity 66.7%; Pred. No. 0.12;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LEAQICQIIEYFGD 17
Db 10 LEKVAEQIIEYFGD 24

RESULT 6
Q87S15 PRELIMINARY; PRT; 390 AA.
ID Q87S15
AC Q87S15;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE A122034P.
GN LA OR CG10922.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Asghyani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA Guerge R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phuanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Cejnkner S.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY075257; AAL68124.1; -.
DR FLYBase; FBgn0011638; la.
DR GO: GO:0008098; F:5S rRNA primary transcript binding; IDA.
DR GO: GO:0003723; F:RNA binding; NAS.
DR InterPro: IPR002344; Lupus_La.
DR InterPro: IPR006630; Lupus_La_dom.
DR InterPro: IPR00504; RNA_rec_mot.
DR Pfam; PF05383; la; 1.

```

```

DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PSS0102; RRM; 1.
DR PROSITE; PSS0030; RRM_NRP_1; 1.
SQ SEQUENCE 390 AA; 44869 MW; 797FDE26B903C909 CRC64;

Query Match 54.2%; Score 52; DB 5; Length 390;
Best Local Similarity 64.3%; Pred. No. 2.5;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 EAQICQIIEYFGD 17
Db 51 ERAIRQVEYFGD 64

RESULT 7
Q81MP9 PRELIMINARY; PRT; 481 AA.
ID Q81MP9
AC Q81MP9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OSJNB4001114.12.
OS Oryza sativa (Japanese cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriactoidae; Oryzaceae; Oryza.
OX NCBI_Taxid=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RA Sasaki C., Henry D., Oates R., Simmons J.;
RT "rice Genomic Sequence";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC105730; AAM51838.1; -.
DR Gramene; Q81MP9; -.
DR InterPro: IPR006630; Lupus_La_dom.
DR Pfam; PF05383; la; 1.
DR SMART; SM00715; LA; 1.
DR Hypothetical protein.
SQ SEQUENCE 481 AA; 49959 MW; 6EA2BED31191362 CRC64;

Query Match 54.2%; Score 52; DB 10; Length 481;
Best Local Similarity 60.0%; Pred. No. 3.1;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LEAQICQIIEYFGD 17
Db 307 LRRAELKQIIEYFGD 321

RESULT 8
Q8NTN8 PRELIMINARY; PRT; 381 AA.
ID Q8NTN8
AC Q8NTN8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Capsular polysaccharide synthesis enzyme Cap8P.
GN CAP8P OR MW0139.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_Taxid=196620;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwano N., Asano K., Naito T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-

```

```

DR GO; GO:0005634; C.nucleus; IDA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR Pfam; PF00076; rtm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
SQ SEQUENCE 381 AA; 43891 MW; 2E2DEF1452C0F0E9 CRC64;

Query Match
Best Local Similarity 86.5%; Score 83; DB 11; Length 381;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALEAQCQIQIEYFGDF 18
DB 12 ALEAKICQIEYFGDF 28

RESULT 2
Q8BTU4 PRELIMINARY; PRT; 415 AA.
AC Q8BTU4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Sjoegren syndrome antigen B.
GN SSB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANCOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK086677; BAC040498.1; -.
DR MGD; MGI:98423; Ssb.
DR GO; GO:0005634; C.nucleus; IDA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR Pfam; PF00076; rtm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 415 AA; 47657 MW; A7545C7686AC8363 CRC64;

Query Match
Best Local Similarity 86.5%; Score 83; DB 11; Length 415;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALEAQCQIQIEYFGDF 18
DB 12 ALEAKICQIEYFGDF 28

RESULT 3
Q8QHT5 PRELIMINARY; PRT; 206 AA.
AC Q8QHT5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Untranslated region binding-protein.

```

```

GN UBP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopteria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA L'Ecuver T.J., Fang H.-L.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF467897; AAL76269.1; -.
DR GO; GO:0005634; C.nucleus; IEA.
DR GO; GO:0003723; P-RNA binding; IEA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR Pfam; PF00076; rtm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 206 AA; 23992 MW; 965B62F7DFE9B90E9 CRC64;

Query Match
Best Local Similarity 72.9%; Score 70; DB 13; Length 206;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LEAQICQIQIEYFGD 17
DB 13 LESKICQIEYFGN 27

RESULT 4
Q7ZTK2 PRELIMINARY; PRT; 427 AA.
ID Q7ZTK2;
AC Q7ZTK2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to lupus LA protein homolog B.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC046654; AAH46654.1; -.
DR GO; GO:0005634; C.nucleus; IEA.
DR GO; GO:0003723; P-RNA binding; IEA.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR006630; Lupus_La_dom.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rtm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 427 AA; 48996 MW; 1E7CD82D8A9C69A CRC64;

Query Match
Best Local Similarity 70.8%; Score 68; DB 13; Length 427;
Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LEAQICQIQIEYFGD 17
DB 12 LPTKICQIEYFGD 26

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:29:01 / Search time 41.229 Seconds
(without alignments)
137.751 Million cell updates/sec

Title: US-09-836-073-2
Perfect score: 96
Sequence: 1 AALEAQICQIIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SPTREMBL.25.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83	86.5	381	11 Q9CYB9	Q9CYB9 mus musculus
2	83	86.5	415	11 Q8BTU4	Q8BTU4 mus musculus
3	70	72.9	206	13 Q8QHT5	Q8QHT5 gallus galli
4	68	70.8	427	13 Q7ZTK2	Q7ZTK2 xenopus lae
5	60	62.5	401	13 Q7ZTK10	Q7ZTK10 xenopus lae
6	52	54.2	390	5 Q8T8V5	Q8T8V5 drosophila
7	52	54.2	481	10 Q8LMP9	Q8LMP9 oryza sativ
8	50	52.1	381	16 Q8NTN8	Q8NTN8 staphylococ
9	50	52.1	391	2 P72382	P72382 staphylococ
10	50	52.1	391	2 P95709	P95709 staphylococ
11	50	52.1	391	16 Q9X57	Q9X57 staphylococ
12	47	49.0	411	10 Q9FJ36	Q9FJ36 arabidopsis
13	47	49.0	422	10 Q94A38	Q94A38 arabidopsis
14	47	49.0	478	16 Q87G58	Q87G58 vibrio para
15	45.5	47.4	928	10 Q9L02	Q9L02 oryza sativ
16	45	46.9	396	5 Q01806	Q01806 caenorhabdi

17	45	46.9	688	5 Q9YBE6	Q9YBE6 drosophila
18	45	46.9	826	10 Q940X9	Q940X9 arabidopsis
19	44	45.8	389	10 Q8S0T8	Q8S0T8 oryza sativ
20	44	45.8	506	10 Q9M143	Q9M143 arabidopsis
21	44	45.8	898	10 Q8LRF3	Q8LRF3 oryza sativ
22	44	45.8	1126	5 Q8EPR0	Q8EPR0 cryptospori
23	44	45.8	1126	5 Q7YXV0	Q7YXV0 cryptospori
24	44	45.8	1126	5 Q44678	Q44678 caenorhabdi
25	43	44.8	111	17 Q978R4	Q978R4 thermoplasma
26	43	44.8	119	10 Q9ZP89	Q9ZP89 neurotheca
27	43	44.8	181	16 Q8Z279	Q8Z279 enterococcu
28	43	44.8	271	16 Q8Z2F4	Q8Z2F4 streptococce
29	43	44.8	355	10 Q8RYB4	Q8RYB4 phytophthor
30	43	44.8	506	10 Q84TQ4	Q84TQ4 amni majus
31	43	44.8	545	10 Q80567	Q80567 arabidopsis
32	42	43.8	151	5 Q86WR7	Q86WR7 plasmodium
33	42	43.8	159	16 Q97LL1	Q97LL1 clostridium
34	42	43.8	186	11 Q8YDM3	Q8YDM3 mus musculu
35	42	43.8	214	11 Q9CS12	Q9CS12 mus musculu
36	42	43.8	219	3 Q9P8N6	Q9P8N6 cochlidiolu
37	42	43.8	328	4 Q8MYR3	Q8MYR3 drosophila
38	42	43.8	269	5 Q8MYR2	Q8MYR2 homo sapien
39	42	43.8	342	11 Q9CJ05	Q9CJ05 mus musculu
40	42	43.8	482	16 Q8YCX0	Q8YCX0 bruceella su
41	42	43.8	482	16 Q8G0Z3	Q8G0Z3 bruceella su
42	42	43.8	575	4 Q8NAP2	Q8NAP2 homo sapien
43	42	43.8	669	16 Q8XRQ1	Q8XRQ1 ralsstonia s
44	42	43.8	965	5 Q9VAM5	Q9VAM5 drosophila
45	42	43.8	1064	16 Q8FHH0	Q8FHH0 corynebacte

ALIGNMENTS

RESULT 1
Q9CYB9 PRELIMINARY; PRT; 381 AA.
ID Q9CYB9
AC Q9CYB9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sjogren syndrome antigen B.
GN SSB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Iehi Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Niole G., Quackenbush J.,
RA Schirml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Maehio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weltz C., Whitlaker C., Wilming L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA "Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK017822; BAB30957.1; -;
DR MGD; MGI:98423; SSB.


```

RP VT12.
RX MEDLINE:20346955; PubMed:1088666;
RT Baasham D.C., Sanderfoot A.A., Kovalева V., Zheng H., Rakhel N.V.;
RT "AtVPS45 complex formation at the trans-Golgi network.";
RL Mol. Biol. Cell 11:2251-2265(2000).
CC -1- FUNCTION: Involved in the protein transport to the vacuole,
CC probably at the level of vesicle fusion at the trans-Golgi network
CC (TGN) and not in transport from the TGN to the prevacuolar
CC compartment. Binds syntaxins.
CC -1- SUBUNIT: Interacts with both SYP41 or SYP42 and VT12, but in
CC different domains of the trans-Golgi network. Does not interact on
CC the prevacuolar compartment with VT11, SYP21 or SYP22, or on the
CC cis-Golgi with SYP31.
CC -1- SUBCELLULAR LOCATION: Peripheral membrane protein. Binds to trans-
CC Golgi network membranes through interaction with other proteins.
CC -1- TISSUE SPECIFICITY: Highly expressed in roots, lower expression in
CC leaves, stems and flowers.
CC -1- SIMILARITY: Belongs to the STXBP/UNC-18/SEC1 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF036234; AAC39472.1; -
CC EMBL; AC004260; AAC34344.1; -
CC EMBL; AY050370; AAC91388.1; -
CC EMBL; AY101517; AAM26638.1; -
CC PIR; T00445; T00445.
CC PIR; T52056; T52056.
CC InterPro; IPR001619; Sec1-like.
CC Pfam; PF00995; Sec1; 1.
CC Transport; Protein transport; T>I (IN RPF. 1).
CC CONFLICT 362 362 64942 MW; 844F24C9A21E9C1D CRC64;
CC SEQUENCE 569 AA; 64942 MW; 844F24C9A21E9C1D CRC64;
CC -----
Query Match 44.8%; Score 43; DB 1; Length 569;
Best Local Similarity 40.0%; Pred. No. 9.8;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 4 EAQICQQLIYFEGD 18
DB 116 EQEYVQVQVQYVADF 130
-----
RESULT 15
SRO9_YEAST STANDARD; PRT; 466 AA.
AC P25567;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SRO9 protein.
GN SRO9 OR YCL037C OR YCL37C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxId=4932;
RN 11;
RP SEQUENCE FROM N.A.
RA Delgado M., Estebe M., Navas L.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN 12;
RP CHARACTERIZATION.
RX MEDLINE=98043396; PubMed=9383048;
RT "SRO9, a multicopy suppressor of the bud growth defect in the
RT Saccharomyces cerevisiae RHO3-deficient cells, shows strong genetic
RT interactions with tropomyosin genes, suggesting its role in
RT organization of the actin cytoskeleton."

```

```

RL Genetics 147:1003-1016(1997).
CC -1- FUNCTION: May overlap in function with tropomyosin and may be
CC involved in organization of actin filaments. Acts as a multicopy
CC suppressor of RHO3.
CC -1- SIMILARITY: Some, to yeast SLP1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X59720; CAA42379.1; -
CC PIR; S19365; S19365.
CC GenOnline; 138873; -
CC SGD; S0000542; SRO9.
CC GO; GO:0005844; C:polysome; IDA.
CC GO; GO:0003723; F:RNA binding; IDA.
CC GO; GO:0006412; P:protein biosynthesis; IPI.
CC InterPro; IPR006630; Lupa_La_dom.
CC Pfam; PF05383; La; 1.
CC SMART; SM00715; LA; 1.
CC DOMAIN 200 203 POLY-GLN.
CC DOMAIN 254 258 POLY-ASN.
CC SEQUENCE 466 AA; 51789 MW; AD8BB0895545117C CRC64;
CC -----
Query Match 42.7%; Score 41; DB 1; Length 466;
Best Local Similarity 63.6%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 7 ICQQLIYFEGD 17
DB 297 IARQLEIYFSE 307
-----
Search completed: September 10, 2004, 17:53:02
Job time : 10.2402 secs

```

RP VARIANT GALACTOSEMIA THR-28.
 RX MEDLINE=99452591; PubMed=10521295;
 RA Kalydjieva L., Perez-Lezaun A., Angelicheva D., Onengut S., Dye D.,
 Rossard N.U., Jordanova A., Savov A., Yanakiev P., Kremensky I.,
 Radeva B., Hallmayer J., Markov A., Nedkova V., Tounnev I., Aneva L.,
 RA Gitzelmann R.,
 RT "A founder mutation in the GALT gene is responsible for galactokinase
 deficiency in Roma (Gypsies)."
 RL Am. J. Hum. Genet. 65:1299-1307(1999).
 RN [6]
 RP VARIANT GALACTOSEMIA VAL-198.
 RX MEDLINE=21152290; PubMed=11211902;
 RA Okano Y., Asada M., Fujimoto A., Ohtake A., Murayama K., Hsiao K.-J.,
 RA Choeh K., Yang Y., Cao Q., Reichardt J.K.V., Nihira S., Imamura T.,
 RA Yamano T.,
 RT "A genetic factor for age-related cataract: identification and
 characterization of a novel galactokinase variant, 'Osaka,' in
 RT Asians.";
 RL Am. J. Hum. Genet. 68:1036-1042(2001).
 CC -1- FUNCTION: Major enzyme for galactose metabolism.
 CC -1- CATALYTIC ACTIVITY: ATP + D-galactose = ADP + D-galactose 1-
 CC phosphate.
 CC -1- PATHWAY: Galactose metabolism; first step.
 CC -1- DISEASE: Defects in GALK1 are the cause of galactosemia II
 CC [MIM:230200], an autosomal recessive deficiency characterized by
 CC congenital cataracts during infancy and presenile cataracts in the
 CC adult population. The cataracts are secondary to accumulation of
 CC galactitol in the lenses.
 CC -1- SIMILARITY: Belongs to the GMP kinase family. GALK subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U26401; AAA96147.1; -;
 CC EMBL: L76927; AAB51607.1; -;
 CC EMBL: BC001166; AAH01166.1; -;
 CC Genew: HGNC:4118; GALK1.
 CC MIM: 60413; -;
 CC MIM: 230200; -;
 CC GO: GO:0005737; C:cytoplasm; TAS.
 CC GO: GO:0004335; F:galactokinase activity; TAS.
 CC GO: GO:0006012; P:galactose metabolism; TAS.
 CC InterPro: IPR001705; Galactokinase.
 CC InterPro: IPR006204; GMP kinase.
 CC InterPro: IPR006203; GMPkinase ATP.
 CC InterPro: IPR006206; Mv_galkinase.
 CC Pfam: PF00288; GMP_kinases; 1.
 CC PRINTS: PR00473; GALTOKINASE.
 CC PRINTS: PR00960; LMBPROTEIN.
 CC PRINTS: PR00959; MEVGAALKINASE.
 CC TIGRfam: TIGR00131; gal_kin; 1.
 CC PROSITE: PS00106; GALACTOKINASE; 1.
 CC PROSITE: PS00627; GMP_KINASES ATP; 1.
 CC Transferrase: Kinase; Galactose metabolism; ATP-binding;
 CC Disease mutation.
 FT NP_BIND 134 144 ATP (POTENTIAL).
 FT VARIANT 28 28 P->T (in galactosemia II).
 FT VARIANT 32 32 V->M (in galactosemia II).
 FT VARIANT 198 198 A->V (in galactosemia II; mild
 FT deficiency; Osaka).
 FT SEQUENCE 392 AA; 42272 MW; 8D7CF8FDB0B4718 CRC64;
 SQ

Query Match 45.8%; Score 44; DB 1; Length 392;
 Best Local Similarity 57.1%; Pred. NO. 4.3;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 Qy 2 ALEAGICQIEYF 15
 Db 164 AARQVCCQAEHSF 177
 RESULT 14
 VP45_ARATH STANDARD; PRT; 569 AA.
 ID VP45_ARATH
 AC 049018; 080650;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Vacuolar protein-sorting protein 45 homolog (AtVP45).
 GN VP45 OR ATIG7140 OR T14N5.2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=98289086; PubMed=9625693;
 RA Bassham D.C., Raikhel N.V.,
 RT "An Arabidopsis VP45p homolog implicated in protein transport to the
 RT vacuole.";
 RL Plant Physiol. 117:407-415(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altafi H., Arujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chao A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Khan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marzalli A.,
 RA Millescher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pail G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salberg S.L., Schwartz J.R., Shin P., Southwick A.M.,
 RA Sun H., Tallon L.V., Tambunga G., Toriumi M.U., Town C.D., Walker M.,
 RA Utecherback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:816-820(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=22954850; PubMed=14593172;
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
 RA Scott-Downing A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Matrand M., Quach H.L., Tripp M., Chang S.H., Lee J.M., Toriumi M.J.,
 RA Chan M.M., Tang C.C., Onders C.S., Deng J.M., Akiyama K., Anarfi Y.,
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjai M., Hansen N.F.,
 RA Hayashizaki Y., Johnson-Hopson C., Heun V.W., Iida K., Karnes M.,
 RA Khan S., Koseme E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamiya A., Meyers E., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
 RA "Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,
 RT "Empirical analysis of transcriptional activity in the Arabidopsis
 RT genome.";
 RL Science 302:842-846(2003).
 RN [4]
 INTERACTIONS WITH SYP1, SYP2, SYP3, SYP4, SYP5, SYP6, SYP7, SYP8, SYP9, SYP10, SYP11 AND

SQ SEQUENCE 487 AA; 54571 MW; 7D4E0E7760A0AF7 CRC64;
 Query Match 46.9%; Score 45; DB 1; Length 487;
 Best Local Similarity 69.2%; Pred. No. 3.7;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 ALEAQCQOIEYV 14
 Db 228 ALBAERCRQIEAY 240
 RESULT 12
 GALT_CANFA STANDARD; PRT; 392 AA.
 ID GALT_CANFA
 AC 09GK4;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Galactokinase (EC 2.7.1.6) (Galactose kinase).
 GN GALK1.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sidjanin D.J.;
 RT "Identification and analysis of canine galactokinase (GALK1) cDNA";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22508185; PubMed=12620391;
 RA Sidjanin D.J., Miller B., Kijas J., McElwee J., Pillardy J., Malek J.,
 RA Pai G., Feldblyum T., Fraser C., Acland G., Aguirre G.;
 RT "Radiation hybrid map, physical map, and low-pass genomic sequence of
 the canine prcd region on CPA9 and comparative mapping with the
 RT syntenic region on human chromosome 17";
 RL Genomics 81:138-148(2003);
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Graves K.T., Emis R.B.;
 RT "Exon scan of the canine galactokinase (GALK1) gene in dog breeds
 RL affected with juvenile cataract";
 Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Major enzyme for galactose metabolism.
 CC -1- CATALYTIC ACTIVITY: ATP + D-galactose = ADP + D-galactose 1-
 CC phosphate.
 CC -1- PATHWAY: Galactose metabolism; first step.
 CC -1- SIMILARITY: Belongs to the GHMP kinase family. Galk subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF213513; AAC34832.1; -
 CC EMBL; AF454963; AA015527.1; -
 CC EMBL; AY267338; AAP31026.1; -
 DR EMBL; AF213513; AAC34832.1; -
 DR EMBL; AF454963; AA015527.1; -
 DR EMBL; AY267338; AAP31026.1; -
 DR InterPro; IPR000705; Galactokinase.
 DR InterPro; IPR001174; GALK_kinase.
 DR InterPro; IPR006204; GHMP_kinase.
 DR InterPro; IPR006203; GHMP_kinase ATP.
 DR InterPro; IPR006206; Mv_galk_kinase.
 DR Pfam; PF00288; GHMP_kinases.1.
 DR PRINTS; PR00473; GALTCTOKINASE.
 DR PRINTS; PR00960; LMPPROTEIN.
 DR PRINTS; PR00959; MEVGALKINASE.
 DR TIGRFAMs; TIGR00131; gal_kin.1.
 DR PROSITE; PS00106; GALACTOKINASE; 1.
 DR PROSITE; PS00627; GHMP_KINASES_ATP; 1.
 DR

KW Transferase; Kinase; Galactose metabolism; ATP-binding.
 FT NP_BIND 134 144 ATP (POTENTIAL).
 SQ SEQUENCE 392 AA; 42120 MW; 8FD5BCF3AB4386BF CRC64;
 Query Match 45.8%; Score 44; DB 1; Length 392;
 Best Local Similarity 57.1%; Pred. No. 4.3;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 2 ALEAQCQOIEYV 15
 Db 164 AARAVQVCOAHSF 177
 RESULT 13
 GALT_HUMAN STANDARD; PRT; 392 AA.
 ID GALT_HUMAN
 AC P51570;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Galactokinase (EC 2.7.1.6) (Galactose kinase).
 GN GALK1 OR GALK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANT GALACTOSEMIA MET-32.
 RX MEDLINE=95400298; PubMed=7670469;
 RA Stambolian D., Al Y., Sidjanin D., Nesburn K., Sathe G., Rosenberg M.,
 RA Bergsma D.J.;
 RT "Cloning of the galactokinase cDNA and identification of mutations in
 RT two families with cataracts";
 Nat. Genet. 10:307-312(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97064967; PubMed=8908517;
 RA Bergsma D.J., Al Y., Skach W.R., Nesburn K., Anicia E.,
 RA van Horn S., Stambolian D.;
 RT "Fine structure of the human galactokinase GALK1 gene";
 RL Genome Res. 6:980-985(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner G.J., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Cabaavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usetin T.B., Toshlyuk S., Carinici P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Wuzny D.N., Sodergren E.V., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.T., Skalska U., Smalls D.E.,
 RA Scherch A., Schein J.E., Jones S.J.W., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=95352063; PubMed=7542884;
 RA Al Y., Basu M., Bergsma D.J., Stambolian D.;
 RT "Comparison of the enzymatic activities of human galactokinase GALK1
 RT and a related human galactokinase protein GK2";
 RL Biochem. Biophys. Res. Commun. 212:687-691(1995).
 RN [5]

DR PIR: T43542; T43542.
 DR GenDB_Spomb: SPAC57A10.10c; -
 DR InterPro: IPR002344; Lupus_La.
 DR InterPro: IPR006630; Lupus_La_dcm.
 DR InterPro: IPR000504; RNA_rec_mol.
 DR Pfam: PF05383; La; 1.
 DR Pfam: PF00076; rrm; 1.
 DR PRINTS: PR00302; LUPUSLA.
 DR SMART: SM00715; La; 1.
 DR SMART: SM00360; RRM; 1.
 DR PROSITE: PS0102; RRM; 1.
 DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
 DR RNA-binding; Nuclear protein.
 DR RNA-binding; Nuclear protein.
 FT DOMAIN 154 235 RNA-BINDING (RRM).
 FT CONFICT 188 188 M -> I (IN RRF 1 AND 2).
 SQ SEQUENCE 298 AA; 34616 MW; 6466AB9940B87F4 CRC64;
 Query Match 51.0%; Score 49; DB 1; Length 298;
 Best Local Similarity 50.0%; Pred. No. 0.42;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 QY 4 EAQICQIEYFSD 17
 Db 64 EAEVLKQVEYFSD 77
 RESULT 10
 ID TCMO_PETCR STANDARD; PRT; 506 AA.
 AC Q43033;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Trans-cinnamate 4-monoxygenase (EC 1.14.13.11) (Cinnamic acid
 4-hydroxylase) (C4H) (C4H) (P450C4H) (Cytochrome P450 73).
 GN CYP73A10 OR CYP73.
 OS Petroselinum crispum (Parsley) (Petroselinum hortense).
 OC Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Campanulids; Apiales; Apiaceae; Apioidae; apioid superclade;
 OC Apium clade; Petroselinum.
 OX NCBI_TaxID=4043;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95320184; PubMed=7597051.
 RA Logemann E., Parniske M., Hanhbrock K.;
 RT "Modes of expression and common structural features of the complete
 phenylalanine ammonia-lyase gene family in parsley.",
 Proc. Natl. Acad. Sci. U.S.A. 92:5905-5909(1995).
 RL
 CC -1- FUNCTION: Controls carbon flux to pigments essential for
 pollination or UV protection, to numerous phycoalexins synthesized
 by plants when challenged by pathogens, and to lignins.
 CC -1- CATALYTIC ACTIVITY: Trans-cinnamate + NADPH + O(2) = 4-
 hydroxycinnamate + NADP(+) + H(2)O.
 CC -1- PATHWAY: Lignin biosynthesis.
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
 CC -1- PATHWAY: Phenylpropanoid metabolism, second step.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL: L38898; AAC41660.1; -
 DR PIR: T14907; T14907.
 DR InterPro: IPR001128; Cytochrome_P450.
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 KM Oxidoreductase; Monooxygenase; Heme; NADP.

FT METAL 448 448 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 506 AA; 58047 MW; 32F0EB959D69CF CRC64;
 Query Match 49.0%; Score 47; DB 1; Length 506;
 Best Local Similarity 50.0%; Pred. No. 1.7;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 5 AQICQIEYFSD 18
 Db 215 SRAQSFYHFGDF 228
 RESULT 11
 ID GATB_CHLGV STANDARD; PRT; 487 AA.
 AC 982386;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Aspartyl/glutamyl-L-Asn(Gln) amidotransferase subunit B
 (EC 6.3.5.-) (Asp/Glu-ADT subunit B).
 GN GATB OR CCA00289
 OS Chlamydomonas reinhardtii.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomonadales.
 OX NCBI_TaxID=83557;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GPIC.
 RX MEDLINE=22569155; PubMed=12682364;
 RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
 RA Heidelberg J., Holtzapple E., Khouri H., Fedorova N.B., Carty H.A.,
 RA Mayhew L.A., Haft D.H., Peterson J., Beanan M.J., White O.,
 RA Salzberg S.L., Hest R.-C., McClarty G., Rank R.G., Bayliss P.M.,
 RA Fraser C.M.;
 RT "Genome sequence of Chlamydomonas reinhardtii (Chlamydia psittaci GPIC):
 examining the role of niche-specific genes in the evolution of the
 Chlamydiales.",
 RT Chlamydiales.
 RL Nucleic Acids Res. 31:2134-2147(2003).
 CC -1- FUNCTION: Allows the formation of correctly charged Asn-tRNA(Asn)
 or Gln-tRNA(Gln) through the transamidation of misacylated Asp-
 tRNA(Asn) or Glu-tRNA(Gln) in organisms which lack either or both
 of asparaginyl-tRNA or glutamyl-tRNA synthetases. The reaction
 takes place in the presence of glutamine and ATP through an
 activated phospho-Asp-tRNA(Asn) or phospho-Glu-tRNA(Gln) (By
 similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + L-glutamyl-L-tRNA(Gln) + L-glutamine = ADP
 + phosphate + L-glutamyl-L-tRNA(Gln) + L-glutamate.
 CC -1- CATALYTIC ACTIVITY: ATP + L-asparaginyl-L-tRNA(Asn) + L-glutamine = ADP
 + phosphate + L-asparaginyl-L-tRNA(Asn) + L-glutamate.
 CC -1- SUBUNIT: Heterotrimer of A, B and C subunits (By similarity).
 CC -1- SIMILARITY: Belongs to the gatB/gatC family. GatB subfamily.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AE016995; AAP05038.1; -
 DR TIGR: CCA00289; -
 DR HAMAP: MF 00121; -; 1.
 DR InterPro: IPR004413; GatB.
 DR InterPro: IPR006107; GatB_cent.
 DR InterPro: IPR006075; GatB_N.
 DR InterPro: IPR003789; GatB_Yqey.
 DR Pfam: PF01162; GatB; 1.
 DR Pfam: PF02934; GatB_N; 1.
 DR Pfam: PF02637; GatB_Yqey; 1.
 DR TIGRFAMs: TIGR00133; gatB; 1.
 DR PROSITE: PS01234; GATB; 1.
 KM Protein biosynthesis; Ligase; Complete proteome.

RA Laako P, Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Matteli B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacleb J.M.,
RA Palazzolo M., Paltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector A.C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "the genome sequence of *Drosophila melanogaster*,"
RL Science 287:2185-2195 (2000).
CC -1- FUNCTION: May be involved in transcription termination by RNA
CC polymerase III. Binds RNA and DNA. Binds to precursors of RNA
CC polymerase III transcripts. May play a specialized role during fly
CC development.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- DEVELOPMENTAL STAGE: Expressed throughout embryonic, larval,
CC pupal, and adult development. Expression throughout the embryo is
CC followed by a restricted pattern of mesodermal expression that is
CC later confined to the visceral mesoderm, gonads, gut, and salivary
CC glands.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/commercial>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U07652; AAA20518.1; -.
DR EMBL: L32988; AAA2176.1; -.
DR EMBL: AE003666; AAF53885.1; -.
DR PIR: A53773; A53773.
DR PIR: A53781; A53781.
DR FLYBase: FBGN0011638; La.
DR GO: GO:0008098; F:5S rRNA primary transcript binding; IDA.
DR GO: GO:003723; F:RNA binding; NAS.
DR InterPro: IPR002344; Lupus_La.
DR InterPro: IPR006630; Lupus_La_dom.
DR InterPro: IPR005054; RNA_rec_mot.
DR Pfam: PF05383; La; 1.
DR Pfam: PF00076; rrm; 1.
DR PRINTS: PR00302; LUPUSLA.
DR SMART: SM00715; La; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS50102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; 1.
KW RNA-binding; Nuclear protein; DNA-binding.
KW DOMAIN
FT 149 234
FT CONFLICT 169 169 A -> T (IN REF. 1).
FT CONFLICT 182 183 KH -> NS (IN REF. 1).
FT CONFLICT 283 283 A -> R (IN REF. 1).
FT CONFLICT 329 329 K -> N (IN REF. 1).
SQ SEQUENCE 350 AA; 44884 MW; A8099288B90446A5 CQC64;

Query Match	54.2%	Score 52;	DB 1;	Length 350;
Best Local Similarity	64.3%;	Pred. No. 0.17;		
Matches	9;	Conservative	2;	Mismatches 3;
				Indels 0;
				Gaps 0;

```

OY      4 EAQICQOIEYFEGD 17
         | :|:|:|:|:|
Db      51 ERAIRQVEYFEGD 64

```

RESULT 9

LAHI_SCHPO	STANDARD;	PRT;	298 AA.
ID	LAHI_SCHPO		
AC	P87058; O13362; Q10458;		
DT	15-JUL-1998 (Rel. 36, Created)		
DT	15-FEB-1998 (Rel. 36, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	The protein homolog (La ribonucleoprotein) (La autoantigen homolog).		
GN	SLA1 OR SPACS7A10.10C.		
OS	Schizosaccharomyces pombe (Fission yeast).		
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;		
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;		
OC	Schizosaccharomyces.		
CX	NCBI_TaxID=4896;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=98067398; PubMed=9404894;		
RA	van Horn D.J., Yoo C.J., Xue D., Shi H., Wolin S.L.;		
RT	"The La protein in Schizosaccharomyces pombe: a conserved yet		
RT	dispensable phosphoprotein that functions in tRNA maturation."		
RL	RNA 3:1434-1443(1997).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RU	Uesumi R.R.U.;		
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=972;		
RX	MEDLINE=21848401; PubMed=11859360;		
RA	Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,		
RA	Brooks K., Peart N., Hayles J., Baker S., Basham D., Bowman S.,		
RA	Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,		
RA	Gallins S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,		
RA	Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,		
RA	James J., Jones L., Jones M., Leather S., McDonald S., McLean J.,		
RA	Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,		
RA	Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,		
RA	Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,		
RA	Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,		
RA	Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,		
RA	Woodward J., Voicakeit E., Aert R., Robben J., Grympez B.,		
RA	Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,		
RA	Gabel C., Fuchs M., Fitzc C., Holzer E., Moestl D., Hilbert H.,		
RA	Borzys K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,		
RA	Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,		
RA	Goffeau A., Cadieu E., Dreno S., Gloux S., Lelare V., Mottier S.,		
RA	Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,		
RA	Lucas M., Roemer M., Gallardin C., Tallada V.A., Garzon A., Rhode G.,		
RA	Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,		
RA	Domínguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,		
RA	Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,		
RA	Shpakovski G.V., Uessey D., Barrell B.G., Nurse P.;		
RT	"The genome sequence of Schizosaccharomyces pombe";		
RL	Nature 415:871-880(2002).		
CC	-1- FUNCTION: Binds to the precursors of polymerase III RNAs.		
CC	-1- SUBCELLULAR LOCATION: Nucleus (Probable).		
CC	-1- FUNCTION: Functions in tRNA maturation.		
CC	-1- SIMILARITY: Contains 1 RNA recognition motif (RNM) domain.		
CC	-1- SIMILARITY: TO VERTEBRATE PROTEIN LA.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; AF022949; AAB82145.1; -		
DR	EMBL; AB011371; BA24981.1; -		
DR	EMBL; Z94864; CAB08173.1; -		
DR	PIR; T38937; T38937.		

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; L00993; AAA39415.1; -
 DR EMBL; BC003820; AAA03820.1; -
 DR EMBL; Y07951; CAA69249.1; -
 DR MGD; MGI:98423; Sab.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR InterPro; IPR002344; Lupus_La.dom.
 DR InterPro; IPR006630; Lupus_La.dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS0102; RRM; 1.
 DR PROSITE; PS0030; RRM_RNP_1; 1.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 DR RNA-binding; Nuclear protein; RNA-BINDING (RRM).
 FT DOMAIN 111 187
 SQ SEQUENCE 415 AA; 4756 MW; 2D75197692FDC933 CRC64;

Query Match 86.5%; Score 83; DB 1; Length 415;
 Best Local Similarity 88.2%; Pred. No. 6.2e-07;
 Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALEAGICQIEYYFGDF 18
 Db 12 ALEAKICQIEYYFGDF 28

RESULT 5
 LAB_XENLA STANDARD; PRT; 427 AA.
 AC P28049;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein homolog B (La ribonucleoprotein B) (La autoantigen
 DE homolog B).
 GN LAB1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocyte;
 RX MEDLINE=93287095; PubMed=8510143;
 RA Scherly D., Stutz F., Lin-Marg N., Clarkson S.G.;
 RT "La proteins from Xenopus laevis. cDNA cloning and developmental
 expression.";
 RL J. Mol. Biol. 231:196-204(1993).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' termin of virtually all nascent
 CC polymerase III transcripts (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DEVELOPMENTAL STAGE: Barely detectable in stage I/II oocytes,
 CC accumulate in stage III/IV oocytes, then exhibit a roughly
 CC constant steady state level in mature oocytes, eggs, and early
 CC embryos.
 CC -1- PTM: Phosphorylated (Probable).
 CC -1- MISCELLANEOUS: There are two forms of La, Laa and Lab, in xenopus.
 CC -1- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X68818; CAA48716.1; -
 DR PIR; S33817; S33817.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La.dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS0102; RRM; 1.
 DR PROSITE; PS0030; RRM_RNP_1; FALSE NEG.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 DR RNA-binding; Nuclear protein; RNA-BINDING (RRM).
 FT DOMAIN 110 202
 SQ SEQUENCE 427 AA; 48995 MW; 45F3146F8934A355 CRC64;

Query Match 70.8%; Score 68; DB 1; Length 427;
 Best Local Similarity 73.3%; Pred. No. 0.00028;
 Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LEAGICQIEYYFGD 17
 Db 12 LDTKICQIEYYFGD 26

RESULT 6
 LAa_XENLA STANDARD; PRT; 428 AA.
 AC P28048;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein homolog A (La ribonucleoprotein A) (La autoantigen
 DE homolog A).
 GN LAa1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocyte;
 RX MEDLINE=93287095; PubMed=8510143;
 RA Scherly D., Stutz F., Lin-Marg N., Clarkson S.G.;
 RT "La proteins from Xenopus laevis. cDNA cloning and developmental
 expression.";
 RL J. Mol. Biol. 231:196-204(1993).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' termin of virtually all nascent
 CC polymerase III transcripts (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DEVELOPMENTAL STAGE: Barely detectable in stage I/II oocytes,
 CC accumulate in stage III/IV oocytes, then exhibit a roughly
 CC constant steady state level in mature oocytes, eggs, and early
 CC embryos.
 CC -1- PTM: Phosphorylated (Probable).
 CC -1- MISCELLANEOUS: There are two forms of La, Laa and Lab, in xenopus.
 CC -1- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

DR PROSITE: PS00030; RRM_RNP_1; 1.
 KM Systemic lupus erythematosus; RNA-binding; Phosphorylation;
 KW Nuclear protein.
 FT DOMAIN 111 187 RNA-BINDING (RRM).
 FT MOD RES 366 366 PHOSPHORYLATION (BY CK2).
 SQ SEQUENCE 408 AA; 46837 MW; EC153C15F9187FC4 CRC64;

Query Match 90.6%; Score 87; DB 1; Length 408;
 Best Local Similarity 88.9%; Pred. No. 1.2e-07;
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQCQIQEYFGDF 18
 DB 11 AALEAKICHOIEYFGDF 28

RESULT 3
 ID LA RAT STANDARD; PRT; 415 AA.
 AC P38656;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
 DE homolog)
 GN SSB OR SS-B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93246255; PubMed=7916708;
 RA Samsel I., Troester H., Bartsch H., Schwemle M., Igloi G.L.,
 RA Bachmann M.;
 RT "Isolation of rat cDNA clones coding for the autoantigen SS-B/La:
 RT detection of species-specific variations";
 RL Gene 1261265-268(1993).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' termini of virtually all nascent
 CC polymerase III transcripts. It is associated with precursor forms
 CC of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S,
 CC and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC
 CC EMBL; X67859; CAA48043.1; -.
 DR PIR; JCI494; JCI494.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La_dom.
 DR InterPro; IPR005054; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SMO0215; LA; 1.
 DR SMART; SMO0360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 KW RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 111 187 RNA-BINDING (RRM).
 SQ SEQUENCE 415 AA; 47777 MW; 033FD9CCEA75F98 CRC64;

Query Match 90.6%; Score 87; DB 1; Length 415;
 Best Local Similarity 88.9%; Pred. No. 1.2e-07;
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQCQIQEYFGDF 18
 DB 11 AALEAKICHOIEYFGDF 28

RESULT 4
 ID LA MOUSE STANDARD; PRT; 415 AA.
 AC P32067;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
 DE homolog)
 GN SSB OR SS-B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93203630; PubMed=8454877;
 RA Topfer F., Gordon T., McCluskey J.;
 RT "Characterization of the mouse autoantigen La (SS-B). Identification
 RT of conserved RNA-binding motifs, a putative ATP binding site and
 RT reactivity of recombinant protein with poly(U) and human
 RT autoantibodies";
 RL J. Immunol. 150:3091-3100(1993).
 CC [2]
 CC SEQUENCE FROM N.A.
 CC STRAIN=FVB/N; TISSUE=Mammary gland;
 CC MEDLINE=23388257; PubMed=12477932;
 CC Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 CC Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 CC Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 CC Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
 CC Datchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 CC Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 CC Brownstein M.J., Ueidl T.B., Toshiyuki S., Carninci P., Prange C.,
 CC Rana S.S., Loguelfano N.A., Peters G.J., Abramson R.D., Mullay S.J.,
 CC Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 CC Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,
 CC Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 CC Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 CC Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 CC Blakeley A.C., Grimwood J., Schmutz J., Myers R.M.,
 CC Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 CC Schnerch A., Schein J.E., Jones S.J.W., Maira M.A.,
 CC "Generation and initial analysis of more than 15,000 full-length
 CC human and mouse cDNA sequences";
 CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC
 CC [3]
 CC SEQUENCE OF 1-11 FROM N.A.
 CC Groelz D., Bachmann M.;
 CC submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' termini of virtually all nascent
 CC polymerase III transcripts. It is associated with precursor forms
 CC of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S,
 CC and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: Phosphorylated (By similarity).
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

FT DOMAIN 111 187 RNA-BINDING (RRM).
 SQ SEQUENCE 404 AA; 46534 MW; 4EE20B5C262AD6A1 CRC64;
 Query Match 90.6%; Score 87; DB 1; Length 404;
 Best Local Similarity 88.9%; Pred. No. 1.2e-07;
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALBAQICQOIEYFQDF 18
 |||||
 DB 11 AALBAKICQOIEYFQDF 28

RESULT 2
 ID LA HUMAN STANDARD; PRT; 408 AA.
 AC P05455;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein (Sjogren syndrome type B antigen) (SS-B) (La
 DE ribonucleoprotein) (La autoantigen).
 GN SSB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89202037; PubMed=2468131;
 RA Chan E.K.L., Sullivan K.F., Tan E.M.;
 RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus
 RT sequences for RNA-binding.";
 RL Nucleic Acids Res. 17:2233-2244(1989).
 RN 12
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89053970; PubMed=3192525;
 RA Chambers J.C., Kenan D., Martin B.J., Keene J.D.;
 RT "Genomic structure and amino acid sequence domains of the human La
 RT autoantigen.";
 RL J. Biol. Chem. 263:18043-18051(1988).
 RN 13
 RP SEQUENCE FROM N.A.
 RC TISSUE-Placenta, and Skeletal muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Ziegler B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haile F.,
 RA Dietchenko L., Marzella K., Farmer A.F., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Muliahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sedergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Heilton B., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield J.S.N., Krzywicki M.I., Skalka U., Smillie D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN 14
 RP SEQUENCE OF 54-408 FROM N.A.
 RX MEDLINE=88199081; PubMed=2452201;
 RA Sturge A.D., Peterson M.G., McNeillage L.J., Whittingham S.,
 RA Coppel R.S.;
 RT "Characteristics and epitope mapping of a cloned human autoantigen
 RT La.";
 RL J. Immunol. 140:3212-3216(1988).
 RN 15

RP SEQUENCE OF 54-97 FROM N.A.
 RX MEDLINE=85166283; PubMed=3856888;
 RA Chambers J.C., Keene J.D.;
 RT "Isolation and analysis of cDNA clones expressing human lupus La
 RT antigen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2115-2119(1985).
 RN 16
 RP FUNCTION.
 RX MEDLINE=89251617; PubMed=2470590;
 RA Gottlieb E., Steltz J.A.;
 RT "Function of the mammalian La protein: evidence for its action in
 RT transcription termination by RNA polymerase III.";
 RL EMBO J. 8:851-861(1989).
 RN 17
 RP PHOSPHORYLATION.
 RX MEDLINE=92707017; PubMed=9054510;
 RA Fan H., Sakulich A.L., Goodier J.L., Zhang X., Qin J., Marale R.J.;
 RT "Phosphorylation of the human La antigen on serine 366 can regulate
 RT recycling of RNA polymerase III transcription complexes.";
 RL Cell 88:707-715(1997).
 RN 18
 RP INTERACTION WITH DDIX15.
 RX MEDLINE=22346609; PubMed=12458796;
 RA Foutaux M.A., Kolman M.J.M., Van der Heijden A., De Jong A.S.,
 RA Van Venrooij W.J., Pruijn G.J.M.;
 RT "The human La (SS-B) autoantigen interacts with DDIX15/hPrp43, a
 RT putative DEAD-box RNA helicase.";
 RL RNA 8:1428-1443(2002).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' terminus of virtually all nascent
 CC polymerase III transcripts. It is associated with precursor forms
 CC of RNA polymerase III transcripts including crRNA and 4.5S, 5S, 7S,
 CC and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDIX15.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE
 CC C-TERMINAL PART OF THE PROTEIN.
 CC -1- PTM: The N-terminus is blocked.
 CC -1- DISEASE: Sera from patients with systemic lupus erythematosus
 CC often contain antibodies that react with the normal cellular
 CC La protein as if this antigen was foreign.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X13697; CAA11985.1; -
 CC EMBL; J04205; AAA51885.1; -
 CC EMBL; BC001289; AAH01289.1; -
 CC EMBL; BC020818; AAH20818.1; -
 CC PIR; A31888; A31888.
 CC GeneW; HGNC:11316; SSB.
 CC MIM; 109090; -
 CC GO; GO:0030529; C:ribonucleoprotein complex; TAS.
 CC GO; GO:0003729; F:RNA binding; TAS.
 CC GO; GO:0000049; F:RNA binding; TAS.
 CC GO; GO:0008334; P:histone mRNA metabolism; TAS.
 CC GO; GO:0006400; P:RNA modification; TAS.
 CC InterPro; IPR002344; Lupus La.
 CC InterPro; IPR006330; Lupus La dom.
 CC InterPro; IPR000504; RNA_rec_mot.
 CC Pfam; PF05383; La; 1.
 CC Pfam; PF00076; tm; 1.
 CC PRINTS; PR00302; LUPUSLA.
 CC SMART; SM00715; LA; 1.
 CC SMART; SM00360; RRM; 1.
 CC PROSITE; PS50102; RRM; 1.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:01:41 ; Search time 7.24022 Seconds
(without alignments)
129.452 Million cell updates/sec

Title: US-09-836-073-2

Perfect score: 96
Sequence: 1 AALEAQICQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	90.6	404	1	LA_BOVIN
2	87	90.6	408	1	LA_HUMAN
3	87	90.6	415	1	LA_RAT
4	83	86.5	415	1	LA_MOUSE
5	68	70.8	427	1	LAB_XENLA
6	64	66.7	428	1	LAB_XENLA
7	56	58.3	383	1	LA_AEDAL
8	52	54.2	390	1	LA_DROME
9	49	51.0	298	1	LAH1_SCHPO
10	47	49.0	506	1	TCMO_PETCR
11	45	46.9	487	1	GATB_CHICV
12	44	45.8	392	1	GALI1_CANFA
13	44	45.8	392	1	GALI1_HUMAN
14	43	44.8	569	1	VP45_ARATH
15	41	42.7	446	1	SRO9_YEAST
16	40	41.7	175	1	IPYR_PSEBK
17	40	41.7	176	1	IPYR_HABIN
18	40	41.7	292	1	IPYR_HABIN
19	40	41.7	292	1	IPYR_HABIN
20	40	41.7	292	1	IPYR_HABIN
21	40	41.7	292	1	IPYR_HABIN
22	40	41.7	389	1	EP1G_DAUCA
23	40	41.7	391	1	GALI1_MOUSE
24	40	41.7	505	1	TCMO_ARATH
25	40	41.7	505	1	TCMO_CICAR
26	40	41.7	505	1	TCMO_GLYEC
27	40	41.7	505	1	TCMO_GLYEC
28	40	41.7	505	1	TCMO_HELTU
29	40	41.7	505	1	TCMO_PEA
30	40	41.7	505	1	TCMO_PHAUV
31	40	41.7	506	1	TCMO_ZINEL
32	40	41.7	506	1	TCMO_MEDSA
33	40	41.7	506	1	TCMO_RUTGR
					TCMO_SOYBN

34	40	41.7	646	1	YE14_SCHPO	013869	schizosacch
35	40	41.7	755	1	SECG_RAT	062825	rattus norv
36	40	41.7	756	1	SECG_HUMAN	060645	homo sapien
37	40	41.7	1180	1	CAA4_BACTI	P14640	bacillus th
38	40	41.7	1418	1	CELL_CAREL	P14641	caenorhabdi
39	40	41.7	2032	1	TRPG_CAREL	093971	caenorhabdi
40	40	41.1	397	1	TRPG_NITEU	Q82W12	nitrosomona
41	39	40.6	363	1	LE3A_ASPEU	P87256	aspergillus
42	39	40.6	475	1	VG41_BP74	P04530	bacteriophag
43	39	40.6	482	1	LBP_RABIT	P17454	oryctolagus
44	39	40.6	490	1	IFTA_HUMAN	014879	homo sapien
45	39	40.6	1164	1	RPO2_COMPX	P17474	compx viru

ALIGNMENTS

RESULT 1
LA_BOVIN STANDARD; PRT; 404 AA.
ID LA_BOVIN
AC P10881;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lapsus La protein homolog (La ribonucleoprotein) (La autocatalytic
homolog).
DE SSB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary; PubMed=2468131;
RX MEDLINE=89202037; Chan E.K.L., Sullivan K.F., Tan E.M.;
RT "Ribonucleoprotein SS-8/La belongs to a protein family with consensus
sequences for RNA-binding."
RT Nucleic Acids Res. 17:2233-2244(1989).
RL
CC -1- FUNCTION: La protein plays a role in the transcription of RNA
polymerase III. It is most probably a transcription termination
factor. Binds to the 3' terminus of virtually all nascent
polymerase III transcripts. It is associated with precursor forms
of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S,
and 7-2 RNAs.
CC -1- SUBUNIT: Interacts with DDX15 (by similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE
C-TERMINAL PART OF THE PROTEIN.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; X13698; CAA31986.1; -.
DR PIR; S03849; S03849.
DR InterPro; IPR002344; Lapsus La.
DR InterPro; IPR006630; Lapsus La dom.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF05383; La; 1.
DR Pfam; PF00076; rtm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00715; LA; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_NRP_1; 1.
KW RNA-binding; Nuclear protein; Phosphorylation.

This Page Blank (uspto)

A:Molecule type: DNA
 A:Residues: 1-166 <DAV>
 A:Cross-references: EMBL:AF039051; PIDN:AB94258.1; GSPDB:GN00023; CESP:C14C6.12
 A:Experimental source: strain Bristol N2; clone C14C6
 C:Genetics:
 A:Gene: CESP:C14C6.12
 A:Map position: 5
 A:Introns: 42/1; 156/3

Query Match 44.8%; Score 43; DB 2; Length 166;
 Best Local Similarity 61.5%; Pred. No. 7.8;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 QICQIEYFGDF 18
 :|||
 Db 81 KICNVIEYMTGDF 93

RESULT 15

T00677

hypothetical protein At2g43970 [imported] - Arabidopsis thaliana

N:Alternate names: hypothetical protein F6E13.10

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 23-Mar-2001

C:Accession: T00677; G84872

R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,

submitted to the EMBL Data Library, June 1998

A:Description: Arabidopsis thaliana chromosome II BAC F6E13 genomic sequence.

A:Reference number: Z14180

A:Accession: T00677

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-529 <ROU>

A:Cross-references: EMBL:AC004005; NID:g3212846; PID:g3212854

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Xoo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.

euze, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 402, 761-768, 1999

A:title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: G84872

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-529 <STO>

A:Cross-references: GB:AE002093; NID:g3212854; PID:ACC23405.1; GSPDB:GN00139

C:Genetics:

A:Gene: F6E13.10; At2g43970

A:Map position: 2

A:Introns: 200/3; 228/1; 259/3; 283/3; 310/2; 344/3; 367/3; 421/3

Query Match 44.8%; Score 43; DB 2; Length 529;

Best Local Similarity 58.3%; Pred. No. 27;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 QICQIEYFGDF 17
 :|||
 Db 196 KIVNOVEYFSD 207

Search completed: September 10, 2004, 18:02:25
 Job time : 17.0782 secs

RESULT 10

T13542
 RNA-binding protein La1 homolog - fission yeast (*Schizosaccharomyces pombe*)
 N:Alternate names: La autoantigen; ribonucleoprotein La homolog
 C:Species: *Schizosaccharomyces pombe*
 C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000
 C/Accession: T13542; T13325
 R:Van Horn, D.J.; Ito, C.D.; Xue, D.; Shi, H.; Wolin, S.L.
 RMA 3, 1434-1443, 1997
 A>Title: The La protein in *Schizosaccharomyces pombe*: a conserved yet dispensable phospho
 A:Reference number: Z22560; MUID:98067398; PMID:9404894
 A/Accession: T13542
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-298 <VAN>
 A/Cross-references: EMBL:AF022949; PIDN:AA82145.1
 R:Utsumi, R.
 submitted to the EMBL Data Library, February 1998
 A>Description: Screening of *S. pombe* cDNA library using *E. coli* defective in signal tran
 A:Reference number: Z22428
 A/Accession: T13325
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-298 <UTS>
 A/Cross-references: EMBL:AB011371; PIDN:BA24981.1
 C/Genetics:
 A:Gene: la1
 A:Function:
 C:Description: the binding of the La protein to RNA precursors is required for the endo
 C:Keyword: phosphoprotein; RNA binding

Query Match 51.0%; Score 49; DB 2; Length 298;
 Best Local Similarity 50.0%; Pred. No. 1.4;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 4 EAQICQIIEYFGD 17
 ||:|:|:|:|:|:|
 Db 64 EAELKQIEYFGD 77

RESULT 11

T14907
 trans-cinnamate 4-monooxygenase (EC 1.14.13.11) - parsley
 C:Species: *Petroselinum crispum* (parsley)
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
 C/Accession: T14907
 R:Logemann, E.; Parniske, M.; Hahlbrock, K.
 Proc. Natl. Acad. Sci. U.S.A. 92, 5905-5909, 1995
 A>Title: Modes of expression and common structural features of the complete phenylalanin
 A:Reference number: Z18260; MUID:95320184; PMID:7597051
 A/Accession: T14907
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1506 <LOG>
 A/Cross-references: EMBL:L38896; NID:G903871; PIDN:AAC41660.1; PID:G903872
 A/Genetics:
 A:Gene: CAH
 C:Function:
 A>Description: catalyzes the formation of *p*-coumaric acid from trans-cinnamic acid

Query Match 49.0%; Score 47; DB 2; Length 506;
 Best Local Similarity 50.0%; Pred. No. 5.4;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 5 AQCICQIIEYFGD 18
 :|:|:|:|:|:|:|
 Db 215 SRLAQSEYHFGD 228

RESULT 12

T30953
 hypothetical protein C44E4.4 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000
 C/Accession: T30953
 R:Sammons, L.; Wohldmann, P.; Gillam, B.
 submitted to the EMBL Data Library, August 1999
 A>Description: The sequence of *C. elegans* cosmid C44E4.
 A:Reference number: Z20945
 A/Accession: T30953
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-396 <SAM>
 A/Cross-references: EMBL:AF003140; PIDN:AA54169.1
 A:Experimental source: strain Bristol N2; clone C44E4
 C/Genetics:
 A:Map position: 1
 A:Introns: 45/1; 114/3
 A>Note: C44E4.4
 C/Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

Query Match 46.9%; Score 45; DB 2; Length 396;
 Best Local Similarity 50.0%; Pred. No. 9;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 4 EAQICQIIEYFGD 17
 :|:|:|:|:|:|:|
 Db 13 DQKIKQIEYFGN 26

RESULT 13

F85016
 probable RING zinc finger protein [imported] - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 17-May-2002
 C/Accession: F85016
 R:Anonymous, The European Union *Arabidopsis* Genome Sequencing Consortium, The Cold Sprin
 Nature 402, 769-777, 1999
 A>Title: Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*.
 A:Reference number: AB5001; MUID:20083488; PMID:10617198
 A/Accession: F85016
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-506 <STO>
 A/Cross-references: GB:NC_001268; NID:G7267624; PIDN:CA880936.1; GSPDB:GN00140
 C/Genetics:
 A:Gene: AT4G01270
 A:Map position: 4
 C/Superfamily: *Arabidopsis thaliana* hypothetical protein A_IG002N01.19

Query Match 45.8%; Score 44; DB 2; Length 506;
 Best Local Similarity 66.7%; Pred. No. 17;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 5 AQCICQIIEYFG 16
 :|:|:|:|:|:|:|
 Db 491 SQGCIQIEHYFG 502

RESULT 14

T32701
 hypothetical protein C14C6.12 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 28-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C/Accession: T32701
 R:David, M.; Wohldmann, P.; Bauer, C.; Antoniou, B.
 submitted to the EMBL Data Library, December 1997
 A>Description: The sequence of *C. elegans* cosmid C14C6.
 A:Reference number: Z21210
 A/Accession: T32701
 A>Status: preliminary; translated from GB/EMBL/DBJ

A:Reference number: S03848; MUID:89202037; PMID:2468131
 A:Accession: S03848
 A:Molecule type: mRNA
 A:Residues: 1-408 <CH2>
 A:Cross-references: EMBL:X13697; NID:g36414; PIDD:CAA31985.1; PIDD:g36415
 R:Chambers, J.C.; Keene, J.D.
 Proc. Natl. Acad. Sci. U.S.A. 82, 2115-2119, 1985
 A>Title: Isolation and analysis of cDNA clones expressing human lupus La antigen.
 A:Reference number: A22956; MUID:85166283; PMID:3856888
 A:Accession: A22956
 A:Molecule type: mRNA
 A:Residues: 45-97, 'LK', <CH3>
 A:Cross-references: GB:J04205
 A>Note: This sequence has been revised in reference A31868
 R:Nyman, U.; Ringertz, N.R.; Petersen, I.
 Immunol. Lett. 22, 65-72, 1989
 A>Title: Demonstration of an amino terminal La epitope recognized by human anti-La sera.
 A:Reference number: A61051; MUID:89379261; PMID:2476379
 A:Accession: A61051
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-19, 'E', '21-47' <NTM>
 R:Sturges, A.D.; Peterson, M.G.; McNeillage, L.J.; Whittingham, S.; Coppel, R.L.
 J. Immunol. 140, 3212-3218, 1988
 A>Title: Characteristics and epitope mapping of a cloned human autoantigen La.
 A:Reference number: S11013; MUID:8819081; PMID:2452201
 A:Accession: S11013
 A:Molecule type: mRNA
 A:Residues: 'E', '55-287', 'V', '289-408' <STU>
 A:Cross-references: EMBL:M0328; NID:g337456; PIDD:AAA36577.1; PIDD:g337457
 R:Kohaka, H.; Yamamoto, K.; Fujii, H.; Miura, H.; Miyasaka, N.; Nishio, K.; Miyamoto, J.
 Clin. Invest. 85, 1566-1574, 1990
 A>Title: Fine epitope mapping the human SS-B/La protein: Identification of a distinct La
 A:Reference number: I55553; MUID:90237237; PMID:1692037
 A:Accession: I55553
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 81-107 <RES>
 A:Cross-references: GB:M35261; NID:g338491; PIDD:AAA3652.1; PIDD:g338495
 A:Accession: I70205
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 174-224 <RE2>
 A:Cross-references: GB:M35263; NID:g338492; PIDD:AAA3653.1; PIDD:g338496
 A:Accession: I70206
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 279-342 <RE3>
 A:Cross-references: GB:M35262; NID:g338493; PIDD:AAA3654.1; PIDD:g338497
 C/Comment: This protein associates with a variety of small RNA molecules, most of which
 act as a transcription termination factor.
 C/Keywords: phosphoprotein; RNA binding
 F:112-117/Domain: ribonucleoprotein repeat homology <RRM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:228-408/Domain: phosphorylated #status experimental <PHY>

Query Match 90.6%; Score 87; DB 1; Length 408;
 Best Local Similarity 88.9%; Pred. No. 6.8e-07;
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LALEAQCQIIEYFGDF 18
 |||||:|||||
 Db 11 LALEAKICQIIEYFGDF 28

RESULT 3

JCI494
 ribonucleoprotein La - rat
 N/Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 24-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C/Accession: JCI494; S25145
 R:Semsei, I.; Troester, H.; Bartsch, H.; Schwemme, M.; Igloi, G.L.; Bachmann, M.
 Gene 126, 265-268, 1993
 A>Title: Isolation of rat cDNA clones coding for the autoantigen SS-B/La. Detection of
 A:Reference number: JCI494; MUID:93246255; PMID:7916708
 A:Accession: JCI494
 A:Molecule type: mRNA
 A:Residues: 1-415 <SEM>
 A:Cross-references: GB:X67859; NID:g55778; PIDD:CAA48043.1; PIDD:g55779
 A:Experimental source: liver
 C/Comment: This protein associates with a variety of small RNA molecules, most of which
 act as a transcription termination factor.
 C/Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C/Keywords: phosphoprotein; RNA binding
 F:112-118/Domain: ribonucleoprotein repeat homology <RRM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:227-415/Domain: phosphorylated #status predicted <PHY>

Query Match 90.6%; Score 87; DB 1; Length 415;
 Best Local Similarity 88.9%; Pred. No. 6.9e-07;
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LALEAQCQIIEYFGDF 18
 |||||:|||||
 Db 11 LALEAKICQIIEYFGDF 28

RESULT 4
 S33817
 ribonucleoprotein La.B - African clawed frog
 N/Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C/Species: Xenopus laevis (African clawed frog)
 C/Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C/Accession: S33817; S28544
 R:Scheryl, D.; Stutz, F.; Lin-Marg, N.; Clarkson, S.G.
 J. Mol. Biol. 231, 196-204, 1993
 A>Title: La proteins from Xenopus laevis. cDNA cloning and developmental expression.
 A:Reference number: S33817; MUID:93287095; PMID:8510143
 A:Accession: S33817
 A:Molecule type: mRNA
 A:Residues: 1-427 <SCH>
 A:Cross-references: EMBL:X68818; NID:g64875; PIDD:CAA48716.1; PIDD:g64876
 C/Comment: This protein associates with a variety of small RNA molecules, most of which
 act as a transcription termination factor.
 C/Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C/Keywords: phosphoprotein; RNA binding
 F:111-117/Domain: ribonucleoprotein repeat homology <RRM>
 F:112-117/Region: RNA-binding RNP2 motif
 F:150-157/Region: RNA-binding RNP1 motif
 F:227-427/Domain: phosphorylated #status predicted <PHY>

Query Match 70.8%; Score 68; DB 1; Length 427;
 Best Local Similarity 73.3%; Pred. No. 0.0012;
 Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LALEAQCQIIEYFGDF 17
 |||||:|||||
 Db 12 LPTKICEQIIEYFGDF 26

RESULT 5
 S33818
 ribonucleoprotein La.A - African clawed frog
 N/Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C/Species: Xenopus laevis (African clawed frog)
 C/Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C/Accession: S33818; S26545

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:33:11 ; Search time 14.0782 Seconds
(without alignments)
122.988 Million cell updates/sec

Title: US-09-836-073-2

Perfect score: 96
Sequence: 1 AALEAQICQOIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR.78:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	90.6	404	1 S03849	ribonucleoprotein
2	87	90.6	408	1 A31888	ribonucleoprotein
3	87	90.6	415	1 JC1494	ribonucleoprotein
4	68	70.8	427	1 S33817	ribonucleoprotein
5	64	66.7	428	1 S33818	ribonucleoprotein
6	52	54.2	390	2 A53773	La/SS-B homolog D-
7	52	54.2	390	2 A53781	ribonucleoprotein
8	50	52.1	391	2 H89777	capsular polysacch
9	49	51.0	298	2 T38937	rna binding protei
10	49	51.0	298	2 T38937	RNA-binding protei
11	47	49.0	298	2 T38937	trans-cinnamate 4-
12	45	46.9	396	2 T30953	hypothetical prote
13	44	45.8	506	2 P85016	probable RING zinc
14	43	44.8	166	2 T32701	hypothetical prote
15	43	44.8	529	2 T00677	hypothetical prote
16	43	44.8	569	2 T52056	vacuolar protein s
17	43	44.8	569	2 T00445	vacuolar protein-s
18	42	43.8	159	2 C96967	lactoyglutination 1
19	42	43.8	482	2 A93381	ATP-dependent RNA
20	41	42.7	466	2 S19365	hypothetical prote
21	41	42.7	469	2 C70357	hypothetical prote
22	41	42.7	837	2 B82932	preproteol translo
23	41	42.7	2160	2 T20241	hypothetical prote
24	40	41.7	51	2 G81188	hypothetical prote
25	40	41.7	176	2 G64049	inorganic pyrophos
26	40	41.7	220	2 T08628	hypothetical prote
27	40	41.7	262	2 E75516	conserved hypotet
28	40	41.7	292	1 B65040	yfjB protein - Bsc
29	40	41.7	292	2 A85908	MND kinase (EC 2.7

30	40	41.7	292	2 AF0834	conserved hypotet
31	40	41.7	292	2 E91063	NAD kinase (EC 2.7
32	40	41.7	346	2 AB0942	hypothetical prote
33	40	41.7	389	2 S36638	glycoprotein Epi -
34	40	41.7	467	2 T24889	hypothetical prote
35	40	41.7	503	2 T06522	trans-cinnamate 4-
36	40	41.7	505	2 T09525	trans-cinnamate 4-
37	40	41.7	505	2 JC1458	trans-cinnamate 4-
38	40	41.7	505	2 A47454	trans-cinnamate 4-
39	40	41.7	505	2 A47409	cinnamate-4-hydrox
40	40	41.7	505	2 B90181	Na+/H+ antiporter
41	40	41.7	506	1 S36878	cytochrome P450 -
42	40	41.7	517	2 T10857	trans-cinnamate 4-
43	40	41.7	646	2 T38022	probable GRP-bindi
44	40	41.7	658	2 D96656	hypothetical prote
45	40	41.7	788	2 A71076	hypothetical prote

ALIGNMENTS

RESULT 1

S03849
ribonucleoprotein La - bovine
N/Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
C/Species: Bos primigenius taurus (cattle)
C/Date: 31-Dec-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C/Accession: S03849
R/Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.
Nucleic Acids Res. 17, 2233-2244, 1989
A/Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences
A/Reference number: S03848; MUID:89202037; PMID:2468131
A/Accession: S03849
A/Molecule type: mRNA
A/Residues: 1-404 <CHA>
A/Cross-references: EMBL:X13698; NID:g756; PIDD:CAA31986.1; PIDD:g756
A/Note: part of this sequence was confirmed by protein sequencing
C/Comment: This protein associates with a variety of small RNA molecules, most of which
ay act as a transcription termination factor.
C/Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
C/Keywords: blocked amino end; phosphoprotein; RNA binding
F/112-178/Domain: ribonucleoprotein repeat homology <RNM>
F/113-118/Region: RNA-binding RNP2 motif
F/151-158/Region: RNA-binding RNP1 motif
F/228-404/Domain: phosphorylated #status predicted <PHX>
Query Match 90.6%; Score 87; DB 1; Length 404;
Best Local Similarity 88.9%; Pred. No. 6.7e-07;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
C/Species: Homo sapiens (man)
C/Date: 21-May-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C/Accession: A31888; S03848; A22956; A61051; S11013; I55553; I70205; I70206; A31273
R/Chan, E.K.L.; Kenan, D.; Martin, B.J.; Keene, J.D.
J. Biol. Chem. 265, 18043-18051, 1990
A/Title: Genomic structure and amino acid sequence domains of the human La autoantigen.
A/Reference number: A31888; MUID:89053970; PMID:3192525
A/Accession: A31888
A/Molecule type: mRNA
A/Residues: 1-408 <CHA>
A/Cross-references: GB:U04205; NID:g178666; PIDD:AAA51885.1; PIDD:g178687
R/Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.
Nucleic Acids Res. 17, 2233-2244, 1989
A/Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:01:41; Search time 7.24022 Seconds
(without alignments)
129.452 Million cell updates/sec

Title: US-09-836-073-14

Perfect score: 99
Sequence: 1 AALEAKICHQIEHYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 141681 segs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	100.0	404	LA_BOVIN	P10881 bos taurus
2	99	100.0	408	LA_HUMAN	P08455 homo sapien
3	99	100.0	415	LA_RAT	P38656 rattus norv
4	95	96.0	415	LA_MOUSE	P28067 mus musculu
5	70	70.7	427	LAB_XENLA	P28049 xenopus lae
6	66	66.7	428	LA_XENLA	P28048 xenopus lae
7	55	55.6	383	LA_AEDAL	P26457 aedes albop
8	51	51.5	380	LA_DROME	P40786 drosophila
9	46	46.5	298	LAH1_SCHPO	P87058 schizosacch
10	43	43.4	487	LAH1_SCHPO	P87058 schizosacch
11	43	43.4	506	TCMO_PPCRC	P43033 petroselinu
12	42	42.4	242	GLNO_BACST	P27675 bacillus st
13	42	42.4	482	LBP_RABIT	P17454 cryctolagus
14	41	41.4	251	FOL2_MOUSE	P05665 mus musculu
15	41	41.4	490	IFP4_HUMAN	P14879 homo sapien
16	41	41.4	573	CATT_YEAST	P06115 saccharomyc
17	41	41.4	1164	RPO2_COMPX	P17474 compox virtu
18	41	41.4	1164	RPO2_VACCV	P19798 vaccinia vi
19	41	41.4	1164	RPO2_VARV	P33811 variola vir
20	41	41.4	1220	DPOL_HSVB	P28888 equine heip
21	41	41.4	1221	V143_NPVAC	P24307 autographa
22	40	40.4	264	GRAB_HUMAN	P94863 homo sapien
23	40	40.4	433	THIC_FUSNN	P49160 fusobacteri
24	40	40.4	466	SRO9_YEAST	P25567 saccharomyc
25	40	40.4	602	EXSA_BUCAI	P57550 buchera ap
26	40	40.4	868	PD61_HUMAN	P08455 homo sapien
27	40	40.4	869	PD61_MOUSE	P08455 mus musculu
28	40	40.4	1107	MY1B_MOUSE	P46735 mus musculu
29	40	40.4	1136	MY1B_RAT	P05006 rattus norv
30	39	39.4	482	IFP5_HUMAN	P13355 homo sapien
31	39	39.4	488	RBL_OITIU	P14959 olisthodisc
32	39	39.4	755	SEC6_RAT	P02885 rattus norv
33	39	39.4	756	SEC6_HUMAN	P06645 homo sapien

ALIGNMENTS

34	39	39.4	812	1	PLMN_MOUSE	P20918 mus musculu
35	39	39.4	1048	1	ACQ1_ARATH	P04379 arabidopsis
36	38	38.4	176	1	IPR1_HAEIN	P44529 haemophilus
37	38	38.4	281	1	Y818_PYRAS	P08291 pyrobaculum
38	38	38.4	354	1	ALKB_ARATH	P08898 arabidopsis
39	38	38.4	391	1	CYB_PASTE	P15585 parametium
40	38	38.4	436	1	THIC_CLOPE	P08455 clostridium
41	38	38.4	488	1	RBL_ECTSI	P24313 ectocarpus
42	38	38.4	488	1	RBL_PYLTI	P24651 pylaiaella t
43	38	38.4	591	1	Y875_CHLTR	P08483 chlamydia t
44	38	38.4	902	1	ATMA_SALTY	P36640 salmoneila
45	38	38.4	926	1	ME19_DROME	P24087 drosophila

RESULT 1

LA_BOVIN
ID LA_BOVIN
AC P10881; STANDARD; PRT; 404 AA.
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
DE SSB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Plutary;
RX MEDLINE=89202037; PubMed=2468131;
RA Chan E.K.L., Sullivan K.F., Tan E.M.;
RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences for RNA-binding."
RL Nucleic Acids Res. 17:2233-2244(1989).
CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III transcripts. It is most probably a transcription termination factor. Binds to the 3' terminus of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S, and 7-2 RNAs.
CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE C-TERMINAL PART OF THE PROTEIN.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC EMBL: X13698; CAA1986.1; -.
CC PIR: S03849; S03849.
CC InterPro: IPR002344; Lupus_La.
CC InterPro: IPR006630; Lupus_La.dom.
CC InterPro: IPR000504; RNA_rec_mot.
CC Pfam: PF05383; La; 1.
CC Pfam: PF00076; rrm; 1.
CC PRINTS: PR00302; LUPUSLA.
CC SMART: SM00715; LA; 1.
CC SMART: SM00360; RRM; 1.
CC PROSITE: PS50102; RRM; 1.
CC PROSITE: PS00030; RRM_RNP_1; 1.
CC RNA-binding; Nuclear protein; phosphorylation.

FT DOMAIN 111 187 RNA-BINDING (RRM).
 SQ SEQUENCE 404 AA; 46534 MW; 4EE30B5C262AD6A1 CRC64;
 Query Match 100.0%; Score 99; DB 1; Length 404;
 Best Local Similarity 100.0%; Pred. No. 3,1e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AALEAKICHQIEYYRGDF 18
 |||||
 11 AALEAKICHQIEYYRGDF 28

Db 11 AALEAKICHQIEYYRGDF 28

RESULT 2
 LA HUMAN STANDARD; PRT; 408 AA.
 AC P05455;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein (Sjogren syndrome type B antigen) (SS-B) (La
 DE ribonucleoprotein) (La autoantigen).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=892020377; PubMed=2468131;
 RA Chan E.K.L., Sullivan K.F., Tan E.M.,
 RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus
 RT sequences for RNA-binding".
 RN Nucleic Acids Res. 17:2233-2244 (1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89053970; PubMed=3192525;
 RA Chambers J.C., Kenan D., Martin B.J., Keene J.D.,
 RT "Genomic structure and amino acid sequence domains of the human La
 RT autoantigen".
 RN J. Biol. Chem. 263:18043-18051 (1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Placenta, and Skeletal muscle;
 RC MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carrinci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Pahey J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whittinger M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schultz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [4]
 RP SEQUENCE OF 54-408 FROM N.A.
 RX MEDLINE=88199081; PubMed=2452201;
 RA Sturgees A.D., Peterson M.G., McNeillage L.J., Whittingham S.,
 RA Coppel R.S.,
 RT "Characteristics and epitope mapping of a cloned human autoantigen
 RT La".
 RL J. Immunol. 140:3212-3218 (1988).
 RN [5]

RP SEQUENCE OF 54-97 FROM N.A.
 RX MEDLINE=85166283; PubMed=3856888;
 RA Chambers J.C., Keene J.D.,
 RT "Isolation and analysis of cDNA clones expressing human Lupus La
 RT antigen".
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2115-2119 (1985).
 RN [6]
 RP FUNCTION.
 RX MEDLINE=89251617; PubMed=2470590;
 RA Gottlieb E., Steltz J.A.,
 RT "Function of the mammalian La protein: evidence for its action in
 RT transcription termination by RNA polymerase III".
 RL EMBO J. 8:851-861 (1989).
 RN [7]
 RP PHOSPHORYLATION.
 RX MEDLINE=97207017; PubMed=9054510;
 RA Pan H., Sakulich A.L., Goodier J.L., Zhang X., Qin J., Maraite R.J.,
 RT "Phosphorylation of the human La antigen on serine 366 can regulate
 RT recycling of RNA polymerase III transcription complexes".
 RL Cell 88:707-715 (1997).
 RN [8]
 RP INTERACTION WITH DDX15.
 RX MEDLINE=22346609; PubMed=12458796;
 RA Fouraux M.A., Kolman M.J.M., Van der Heijden A., De Jong A.S.,
 RA Van Veenrooij W.J., Pruijn G.J.M.,
 RT "The human La (SS-B) autoantigen interacts with DDX15/hPrp43, a
 RT putative DEAD-box RNA helicase".
 RL RNA 8:1428-1443 (2002).
 RN [9]
 RP FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' terminal of virtually all nascent
 CC polymerase III transcripts. It is associated with precursor forms
 CC of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S,
 CC and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE
 CC C-TERMINAL PART OF THE PROTEIN.
 CC -1- PM: The N-terminus is blocked.
 CC -1- DISEASE: Sera from patients with systemic lupus erythematosus
 CC often contain antibodies that react with the normal cellular
 CC La protein as if this antigen was foreign.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.jab-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 CC EMBL: X13597; CAA31985.1; -;
 CC EMBL: J04205; AAAS1885.1; -;
 CC EMBL: BC001289; AAH01289.1; -;
 CC EMBL: BC020818; AAH20818.1; -;
 CC F01: A31888; A31888.
 CC GeneW: HGNC:11316; SSB.
 CC MIM: 109090; -;
 CC GO: GO:0030529; C:ribonucleoprotein complex; TAS.
 CC GO: GO:0003729; F:mRNA binding; TAS.
 CC GO: GO:0000049; F:RNA binding; TAS.
 CC GO: GO:0008334; P:histone mRNA metabolism; TAS.
 CC GO: GO:0006400; P:RNA modification; TAS.
 CC InterPro: IPR002344; Lupus_La.
 CC InterPro: IPR006630; Lupus_La_dom.
 CC InterPro: IPR000504; RNA_rec_mot.
 CC Pfam: PF05383; La; 1.
 CC Pfam: PF00076; rim; 1.
 CC PRINTS: PR00302; LUPUSLA.
 CC SMART: SM00715; LA; 1.
 CC SMART: SM00360; RRM; 1.
 CC PROSITE: PS50102; RRM; 1.

DR PROSITE: PS00030; RRM_RNP_1; 1.
 KW Systemic lupus erythematosus; RNA-binding; Phosphorylation;
 KM Nuclear protein.
 FT DOMAIN 111 187 RNA-BINDING (RRM).
 FT MOD_RES 366 366 PHOSPHORYLATION (BY CK2).
 SQ SEQUENCE 408 AA; 46837 MW; EC153C15F9187FC4 CRC64;

Query Match 100.0%; Score 99; DB 1; Length 408;
 Best Local Similarity 100.0%; Pred. No. 3.1e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALEAKICHOIEYFGDF 18
 Db 11 ALEAKICHOIEYFGDF 28

RESULT 3

LA_MOUSE STANDARD; PRT; 415 AA.
 ID LA_MOUSE
 AC P38656;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 GN SSB OR SS-B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCB1_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93246255; PubMed=7916708;
 RA Semsei I., Troester H., Bartsch H., Schwemle M., Igloi G.L., Bachmann M.;
 RT "Isolation of rat cDNA clones coding for the autoantigen SS-B/La: detection of species-specific variations."
 RL Gene 126:265-268(1993).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' terminus of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including crRNA and 4.5S, 5S, 7S, and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: Phosphorylated (By similarity).
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X67859; CAA48043.1; -
 CC PIR: JCI494; JCI494.
 DR InterPro: IPR002344; Lupus_La.
 DR InterPro: IPR006630; Lupus_La_dom.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00383; La; 1.
 DR Pfam: PF00076; rrm; 1.
 DR PRINTS: PR00302; LUPUSLA.
 DR SMART: SM00715; La; 1.
 DR SMART: SM00360; RRM; 1.
 DR PROSITE: PS00102; RRM; 1.
 DR PROSITE: PS00030; RRM_RNP_1; 1.
 KM RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 111 187 RNA-BINDING (RRM).
 SQ SEQUENCE 415 AA; 47777 MW; 033FD9CCE475F98 CRC64;

Query Match 100.0%; Score 99; DB 1; Length 415;
 Best Local Similarity 100.0%; Pred. No. 3.2e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALEAKICHOIEYFGDF 18
 Db 11 ALEAKICHOIEYFGDF 28

Query Match 100.0%; Score 99; DB 1; Length 415;
 Best Local Similarity 100.0%; Pred. No. 3.2e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALEAKICHOIEYFGDF 18
 Db 11 ALEAKICHOIEYFGDF 28

RESULT 4

LA_MOUSE STANDARD; PRT; 415 AA.
 ID LA_MOUSE
 AC P32067;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 GN SSB OR SS-B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCB1_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93203630; PubMed=8454877;
 RA Topfner F., Gordon T., McCluskey J.;
 RT "Characterization of the mouse autoantigen La (SS-B). Identification of conserved RNA-binding motifs, a putative ATP binding site and reactivity of recombinant protein with poly(U) and human autoantibodies."
 RL J. Immunol. 150:3091-3100(1993).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX STRAIN=FVB/N; TISSUE=Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Shapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullen S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., White J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Rodriguez Y.S.N., Krzywinski M.I., Skalska U., Smalys D.E., Schnerker A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN (3)
 RP SEQUENCE OF 1-11 FROM N.A.
 RA Grodz D., Bachmann M.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Binds to the 3' terminus of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including crRNA and 4.5S, 5S, 7S, and 7-2 RNAs.
 CC -1- SUBUNIT: Interacts with DDX15 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: Phosphorylated (By similarity).
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; L00993; AAA9415.1; -
 DR EMBL; BC003820; AAA03820.1; -
 DR EMBL; X07951; CAA69249.1; -
 DR MGD; MGI:98423; Sdb.
 DR GO; GO:0005634; C:nucleus, IDA.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La.dom.
 DR InterPro; IPR00504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; xrm; 1.
 DR PRINTS; PR0302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS0030; RRM_RNP_1; 1.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 DR RNA-binding; Nuclear protein; RNA-BINDING (RRM).
 FT DOMAIN 111 187
 SQ SEQUENCE 415 AA; 47756 MW; 2D75197692FDC933 CRC64;

Query Match 96.0%; Score 95; DB 1; Length 415;
 Best Local Similarity 100.0%; Pred. No. 1.5e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALEAKICHOIEYFGDF 18
 Db 12 ALEAKICHOIEYFGDF 28

RESULT 5
 LAB_XENLA STANDARD; PRT; 427 AA.
 ID LAB_XENLA
 AC P28049;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein homolog B (La ribonucleoprotein B) (La autoantigen
 DE homolog B).
 GN LAB1.
 OS Xenopus laevis (African clawed frog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 CC Xenopodinae; Xenopus.
 CC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocyte;
 RX MEDLINE=93287095; PubMed=8510143;
 RA Scherly D., Stutz F., Lin-Marg N., Clarkson S.G.;
 RT "La proteins from Xenopus laevis. cDNA cloning and developmental
 RT expression.";
 RL J. Mol. Biol. 231:196-204(1993).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' terminus of virtually all nascent
 CC polymerase III transcripts (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DEVELOPMENTAL STAGE: Barely detectable in stage I/II oocytes,
 CC accumulate in stage III/IV oocytes, then exhibit a roughly
 CC constant steady state level in mature oocytes, eggs, and early
 CC embryos.
 CC -1- PTM: Phosphorylated (Probable).
 CC -1- MISCELLANEOUS: There are two forms of La, Laa and Lab, in xenopus.
 CC -1- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X68818; CAA48716.1; -
 DR PIR; S33817; S33817.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La.dom.
 DR InterPro; IPR00504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; xrm; 1.
 DR PRINTS; PR0302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS0030; RRM_RNP_1; FALSE NEG.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 DR RNA-binding; Nuclear protein; RNA-BINDING (RRM).
 FT DOMAIN 110 202
 SQ SEQUENCE 427 AA; 48995 MW; 45F3146F934A355 CRC64;

Query Match 70.7%; Score 70; DB 1; Length 427;
 Best Local Similarity 80.0%; Pred. No. 0.00027;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LEAKICHOIEYFGD 17
 Db 12 LDTKICHOIEYFGD 26

RESULT 6
 LAB_XENLA STANDARD; PRT; 428 AA.
 ID LAB_XENLA
 AC P28048;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lupus La protein homolog A (La ribonucleoprotein A) (La autoantigen
 DE homolog A).
 GN LAB1.
 OS Xenopus laevis (African clawed frog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 CC Xenopodinae; Xenopus.
 CC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocyte;
 RX MEDLINE=93287095; PubMed=8510143;
 RA Scherly D., Stutz F., Lin-Marg N., Clarkson S.G.;
 RT "La proteins from Xenopus laevis. cDNA cloning and developmental
 RT expression.";
 RL J. Mol. Biol. 231:196-204(1993).
 CC -1- FUNCTION: La protein plays a role in the transcription of RNA
 CC polymerase III. It is most probably a transcription termination
 CC factor. Binds to the 3' terminus of virtually all nascent
 CC polymerase III transcripts (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DEVELOPMENTAL STAGE: Barely detectable in stage I/II oocytes,
 CC accumulate in stage III/IV oocytes, then exhibit a roughly
 CC constant steady state level in mature oocytes, eggs, and early
 CC embryos.
 CC -1- PTM: Phosphorylated (Probable).
 CC -1- MISCELLANEOUS: There are two forms of La, Laa and Lab, in xenopus.
 CC -1- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X68817; CAA48715.1; -
 DR PIR: S33818; S33818
 DR InterPro: IPR002344; Lupus_La.
 DR InterPro: IPR006630; Lupus_La_dom.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF05383; LA; 1.
 DR Pfam: PF00076; rrm; 1.
 DR PRINTS: PR00302; LUPUSLA.
 DR SMART: SM00715; LA; 1.
 DR SMART: SM00360; RRM; 1.
 DR PROSITE: PS50102; RRM; 1.
 DR PROSITE: PS00030; RRM_RNP_1; 1.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 DR RNA-binding; Nuclear protein; RNA-BINDING (RRM).
 FT DOMAIN 111 203 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 316 332 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 316 332 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 SQ SEQUENCE 428 AA; 48864 MW; AEB3A38B7D2E3EC3 CRC64;
 Query Match 66.7%; Score 66; DB 1; Length 428;
 Best Local Similarity 78.6%; Pred. No. 0.0013; Mismatches 2; Indels 0; Gaps 0;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 4 EAKICHQIEYFGD 17
 DB 14 DTKICEQIEYFGD 27
 RESULT 7
 LA AEDAL STANDARD; PRT; 383 AA.
 ID LA AEDAL
 AC 026457;
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-FEB-2003 (Rel. 41, Last annotation update)
 DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 OS Aedes albopictus (Forest day mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.
 NC NCB1_TaxID=7160;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=96135233; PubMed=8551578;
 RA Parignon N., Straus J.H.;
 RT "Mosquito homolog of the La autoantigen binds to Sindbis virus RNA.";
 RL J. Virol. 70:1173-1181 (1996).
 CC -1- FUNCTION: May be involved in transcription termination by RNA
 CC polymerase III. Binds RNA and DNA. Binds to the 3' end of the
 CC minus strand of Sindbis virus RNA. This may be significant for
 CC Sindbis virus RNA replication.
 CC -1- SUBCELLULAR LOCATION: Nuclear. Primarily nuclear, but significant
 CC amounts are present in the cytoplasm.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -1- SIMILARITY: TO VERTERATE PROTEIN LA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: S80954; AAB35931.1; -
 DR InterPro: IPR002344; Lupus_La.
 DR InterPro: IPR006630; Lupus_La_dom.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF05383; La; 1.
 DR Pfam: PF00076; rrm; 1.
 DR PRINTS: PR00302; LUPUSLA.

DR SMART: SM00715; LA; 1.
 DR SMART: SM00360; RRM; 1.
 DR PROSITE: PS50102; RRM; 1.
 DR PROSITE: PS00030; RRM_RNP_1; FALSE_NBG.
 DR RNA-binding; Nuclear protein; DNA-BINDING (RRM).
 DR RNA-binding; Nuclear protein; RNA-BINDING (RRM).
 FT DOMAIN 141 228
 FT DOMAIN 141 228
 SQ SEQUENCE 383 AA; 44430 MW; 4E5CC8F21C40F452 CRC64;
 Query Match 55.6%; Score 55; DB 1; Length 383;
 Best Local Similarity 66.7%; Pred. No. 0.085;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 3 LEAKICHQIEYFGD 17
 DB 43 LEASTIRQIEYFGD 57
 RESULT 8
 LA DROME STANDARD; PRT; 390 AA.
 ID LA DROME
 AC P40796; Q24375; Q9V1N2;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 GN LA OR CG10922.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCB1_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=Canton-S; TISSUE=Ovary;
 RX MEDLINE=94309632; PubMed=8035794;
 RA Bai C., Li Z., Tollas P.P.;
 RT "Developmental characterization of a Drosophila RNA-binding protein
 RT homologous to the human systemic lupus erythematosus-associated
 RT La/SS-B autoantigen.";
 RL Mol. Cell. Biol. 14:5123-5129 (1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94309661; PubMed=8035818;
 RA Yoo C.J., Wolin S.L.;
 RT "La proteins from Drosophila melanogaster and Saccharomyces
 RT cerevisiae: a yeast homolog of the La autoantigen is dispensable for
 RT growth.";
 RL Mol. Cell. Biol. 14:5412-5424 (1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=Beckley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Abpayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo C., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milbina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Pacleb U.M., Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., She B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier B., Spradling A.C., Stapleton M., Strong R., Sun E., Svrtkals R., Tecor C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of *Drosophila melanogaster*.";
 Science 287:2185-2195(2000).
 -1- FUNCTION: May be involved in transcription termination by RNA polymerase III. Binds RNA and DNA. Binds to precursors of RNA polymerase III transcripts. May play a specialized role during fly development.
 CC SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DEVELOPMENTAL STAGE: Expressed throughout embryonic, larval, pupal, and adult development. Expression throughout the embryo is followed by a restricted pattern of mesodermal expression that is later confined to the visceral mesoderm, gonade, gut, and salivary glands.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U07652; AAA20518.1; -;
 DR EMBL; L32988; AAA21776.1; -;
 DR EMBL; AE003666; AAF53885.1; -;
 DR PIR; A53773; A53773.
 DR PIR; A53781; A53781.
 DR FlyBase; FBgn0011638; La.
 DR GO; GO:0008099; F:5S rRNA primary transcript binding; IDA.
 DR GO; GO:0003723; F:RNA binding; NAS.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006530; Lupus_La_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00383; La; 1.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 KW RNA-binding; Nuclear protein; DNA-binding.
 FT DOMAIN 149 234
 FT CONFLICT 169 169 A->T (IN REF. 1).
 FT CONFLICT 182 183 KH->NS (IN REF. 1).
 FT CONFLICT 283 283 A->R (IN REF. 1).
 FT CONFLICT 329 329 K->N (IN REF. 1).
 SQ SEQUENCE 390 AA; 44884 MW; A809288B90446A5 CRC64;

Query Match 51.5%; Score 51; DB 1; Length 390;
 Best Local Similarity 64.3%; Pred. No. 0.41;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 EAKICHOLEYFGD 17
 DB 51 ERAIRROVEYTFGD 64

RESULT 9
 LAH1 SCHPO STANDARD; PRT; 298 AA.
 ID LAH1 SCHPO
 AC P87058; Q13352; Q10458;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 GN SLA1 OR SPAC57A10.10C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98067398; PubMed=9404894;
 RA van Horn D.J., Yoo C.U., Xue D., Shi H., Molin S.L.;
 RT "The La protein in *Schizosaccharomyces pombe*: a conserved yet dispensable phosphoprotein that functions in rRNA maturation.";
 RL RNA 3:1434-1443(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Utsunli R.R.U.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21849401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A., Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Feltwell J., Fraser A., Gentles M., Goble A., Hamlin N., Harris A., Hidalgo T., Hodgson G., Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jørgels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volkert G., Aert R., Roben J., Grymonprez B., Walteijens I., Vanstreels E., Rieger M., Schefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M., Egert P., Zimmermann W., Wedler H., Mambutt R., Purnelle B., Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Domagala L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.,
 RA "The genome sequence of *Schizosaccharomyces pombe*.";
 RT Nature 415:871-880(2002).
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: Binds to the precursors of polymerase III RNAs.
 CC Functions in rRNA maturation.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF022949; AAB82145.1; -;
 DR EMBL; AB011371; BAA24981.1; -;
 DR EMBL; Z94864; CAB08173.1; -;
 DR PIR; T38937; T38937.

DR PIR; T43542; T43542.
 DR Genedb_Spomb; SPAC57A10.10c; -.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR006630; Lupus_La.dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF05383; La; 1.
 DR Pfam; PF00076; rtm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00715; LA; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PSS0102; RRM; 1.
 DR PROSITE; PSS00030; RRM_RNR_1; FALSE_NEG.
 DR RNA-binding; Nuclear protein.
 DR DOMAIN 154 236 RNA-BINDING (RRM).
 FT CONFLICT 188 188 M -> I (IN REF. 1 AND 2).
 SQ SEQUENCE 298 AA; 34616 MW; 64E6AB9940B87F4 CRC64;
 Query Match 46.5%; Score 46; DB 1; Length 298;
 Best Local Similarity 50.0%; Pred. No. 2.2;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 Oy 4 EAKICHQIEYFGD 17
 Db 64 EAEVLKQVEFYFD 77
 RESULT 10
 GATB_CHLCV STANDARD; PRT; 487 AA.
 ID GATB_CHLCV
 AC 0823W6;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Aspartyl-glutamyl-tRNA (Asn/Gln) amidotransferase subunit B
 DE (EC 6.3.5.-) (Asp/Glu-Adt subunit B).
 GN GATB OR CCA00289.
 OS Chlamydomonas reinhardtii.
 OC Bacteria; Chlamydiales; Chlamydiales; Chlamydiales.
 OX NCBI_TaxID=83557;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPIC.
 RX MEDLINE=22369155; PubMed=12682364;
 RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
 Heidelberg J., Holtzaple E., Khouri H., Federova N.B., Carty H.A.,
 Umayam L.A., Haft D.H., Peterson J., Beanan M.J., White O.,
 Salzberg S.L., Hsiao R.-C., McClarty G., Rank R.G., Bavolli P.M.,
 Fraser C.M.;
 RT "Genome sequence of Chlamydomonas reinhardtii (Chlamydia psittaci GPIC):
 examining the role of niche-specific genes in the evolution of the
 Chlamydiales";
 RT Chlamydiales.
 RL Nucleic Acids Res. 31(21):2134-2147(2003).
 CC -1- FUNCTION: Allows the formation of correctly charged Asn-tRNA(Asn)
 or Gln-tRNA(Gln) through the transamination of misacylated Asp-
 tRNA(Asn) or Glu-tRNA(Gln) in organisms which lack either or both
 of aspartyl-tRNA or glutamyl-tRNA synthetases. The reaction
 takes place in the presence of glutamine and ATP through an
 activated phospho-Asp-tRNA(Asn) or phospho-Glu-tRNA(Gln) (By
 similarity)
 CC -1- CATALYTIC ACTIVITY: ATP + L-glutamyl-tRNA(Gln) + L-glutamine = ADP
 + phosphate + L-glutamyl-tRNA(Gln) + L-glutamate.
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartyl-tRNA(Asn) + L-glutamine = ADP
 + phosphate + L-aspartyl-tRNA(Asn) + L-glutamate.
 CC -1- SUBUNIT: Heterotrimer of A, B and C subunits (By similarity).
 CC -1- SIMILARITY: Belongs to the gatB/gatC family. GatB subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; AE016995; AAP05038.1; -.
 DR TIGR; CCA00289; -.
 DR HAMAP; MF_00121; -; 1.
 DR InterPro; IPR004413; GatB.
 DR InterPro; IPR006107; GatB_cent.
 DR InterPro; IPR006075; GatB_N.
 DR InterPro; IPR003789; GatB_Ygey.
 DR Pfam; PF01162; GatB; 1.
 DR Pfam; PF02934; GatB_N; 1.
 DR Pfam; PF02637; GatB_Ygey; 1.
 DR TIGRfam; TIGR00133; GatB; 1.
 DR PROSITE; PSS0124; GATB; 1.
 DR Protein biosynthesis; Ligase; Complete proteome.
 KW SEQUENCE 487 AA; 54571 MW; 7DAB0EE7760A0AF7 CRC64;
 SQ SEQUENCE
 Query Match 43.4%; Score 43; DB 1; Length 487;
 Best Local Similarity 69.2%; Pred. No. 12;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Oy 2 ALEAKICHQIEY 14
 Db 228 ALEAERCQIEYAY 240
 RESULT 11
 TCMO_PETCR STANDARD; PRT; 506 AA.
 ID TCMT_PETCR
 AC 043033;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Trans-cinnamate 4-monooxygenase (EC 1.14.13.11) (Cinnamic acid
 4-hydroxylase) (C4H) (P450C4H) (Cytochrome P450 73).
 GN CYP73A10 OR CYP73.
 OS Petroselinum crispum (Parsley) (Petroselinum hortense).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Campanulids; Apiales; Apiaceae; Apioidae; apioid superclade;
 OC Apium clade; Petroselinum.
 OX NCBI_TaxID=4043;
 [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=95320184; PubMed=7597051;
 RA Logemann E., Parniske M., Hahlbrock K.;
 RT "Modes of expression and common structural features of the complete
 RT phenylalanine ammonia-lyase gene family in parsley";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:5905-5909(1995).
 CC -1- FUNCTION: Controls carbon flux to pigments essential for
 CC pollination or UV protection, to numerous phytoalexins synthesized
 CC by plants when challenged by pathogens, and to lignins.
 CC -1- CATALYTIC ACTIVITY: Trans-cinnamate + NADPH + O(2) = 4-
 CC hydroxycinnamate + NADP(+) + H(2)O.
 CC -1- PATHWAY: Lignin biosynthesis.
 CC -1- PATHWAY: Phenylpropanoid metabolism; second step.
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

SQ SEQUENCE 506 AA; 58047 MW; 32F00EE959D69CCF CRC64;
 Query Match 43.4%; Score 43; DB 1; Length 506;
 Best Local Similarity 42.9%; Pred. No. 12;
 Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 AKICHIQIEYFGDF 18
 : : : : :
 Db 215 SRLAQSFYHFGDF 228

RESULT 12
 GLNQ_BACST STANDARD; PRT; 242 AA.
 AC P27675;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glutamine transport ATP-binding protein glnQ.
 GN GLNQ.
 OS *Bacillus stearothermophilus*.
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
 CX NCBI_TaxID=1422;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NUB36;
 RX MEDLINE=91310597; PubMed=1856180;
 RA Wu L., Welker N.E.;
 RT "Cloning and characterization of a glutamine transport operon of
Bacillus stearothermophilus NUB36: effect of temperature on
 regulation of transcription.";
 RL J. Bacteriol. 173:4877-4888 (1991).
 CC -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
 FOR GLUTAMINE. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO THE
 TRANSPORT SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Membrane-associated (Potential).
 CC -1- INDUCTION: By lack of glutamine.
 CC -1- SIMILARITY: Belongs to the ABC transporter family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL; M61017; AAA22483.1; -.
 DR PIR; A42478; A42478.
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR003439; ABC_transporter.
 DR Pfam; PF00005; ABC_tran.1
 DR ProDom; PD000006; ABC_transporter.1.
 DR SMART; SM00382; AAA.1.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
 DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
 KW Amino-acid transport; Transport; Membrane; ATP-binding.
 FT NP BIND 34 41
 FT SEQUENCE 242 AA; 27436 MW; 102B1C5E332F31C8 CRC64;

Query Match 42.4%; Score 42; DB 1; Length 242;
 Best Local Similarity 60.0%; Pred. No. 8.4;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 HOIEYFGDF 18
 : : : : :
 Db 5 HOVNKYGDF 14

RESULT 13
 LBP_RABIT STANDARD; PRT; 482 AA.
 ID LBP_RABIT
 AC P17454;

DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Lipopolysaccharide-binding protein precursor (LBP).
 GN LBP.
 OS *Oryctolagus cuniculus* (Rabbit).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Euteria; Lagomorpha; Leporidae; Oryctolagus.
 CX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=90385281; PubMed=2402637;
 RA Schumann R.R., Leong S.R., Flagg G.W., Gray P.W., Wright S.D.,
 RA Mathison J.C., Tobias P.S., Ulevitch R.J.;
 RT "Structure and function of lipopolysaccharide binding protein.";
 RL Science 249:1429-1431 (1990).
 RN [2]
 RP SEQUENCE OF 27-66.
 RC TISSUE=Serum;
 RX MEDLINE=86306528; PubMed=2427635;
 RA Tobias P.S., Soldau K., Ulevitch R.J.;
 RT "Isolation of a lipopolysaccharide-binding acute phase reactant from
 rabbit serum.";
 RL J. Exp. Med. 164:777-793 (1986).
 CC -1- FUNCTION: Binds to the lipid moiety of bacterial
 lipopolysaccharides (LPS), a glycolipid present in the outer
 membrane of all Gram-negative bacteria. The LBP/LPS complex seems
 to interact with the CD14 receptor.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to the BPI/LBP/Plunc superfamily. BPI/LBP
 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL; M35534; AAA99235.1; -.
 DR PIR; B35843; B35843.
 DR HSSP; P17213; LBPI.
 DR InterPro; IPR001124; LBP_BPI_CETP.
 DR Pfam; PF01273; LBP_BPI_CETP.1.
 DR Pfam; PF02886; LBP_BPI_CETP_C.1.
 DR SMART; SM00328; BPI.1.
 DR SMART; SM00329; BPI2.1.
 DR PROSITE; PS00400; LBP_BPI_CETP.1.
 KW Lipid transport; Antibiotic; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 482
 FT CARBOHYD 301 301
 FT CARBOHYD 351 351
 FT CARBOHYD 387 387
 FT CONFLICT 57 57
 FT CONFLICT 63 63
 FT SEQUENCE 482 AA; 54001 MW; 628A6E0A47200C2 CRC64;

Query Match 42.4%; Score 42; DB 1; Length 482;
 Best Local Similarity 80.0%; Pred. No. 17;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LEAKICHOIE 12
 : : : : :
 Db 194 LESKICHOIE 203

RESULT 14
 FOL2_MOUSE STANDARD; PRT; 251 AA.
 ID FOL2_MOUSE
 AC Q05685;
 DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Folate receptor beta precursor (FR-beta) (Folate receptor 2) (Folate-binding protein 2).
 GN FOLR2 OR FOLBP2 OR FBP2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RX MEDLINE=91373339; PubMed=1894617;
 RA Brigle K.E., Westin E.H., Houghton M.T., Goldman I.D.;
 RT "Characterization of two cDNAs encoding folate-binding proteins from RT 11210 murine leukemia cells. Increased expression associated with a genomic rearrangement.";
 RL J. Biol. Chem. 266:17243-17249(1991).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94140851; PubMed=8307991;
 RA Brigle K.E., Seltzer R.L., Westin E.H., Goldman I.D.;
 RT "Increased expression and genomic organization of a folate-binding protein homologous to the human placental isoform in 11210 murine leukemia cell lines with a defective reduced folate carrier.";
 RL J. Biol. Chem. 269:4267-4272(1994).
 -1- FUNCTION: Binds to folate and reduced folic acid derivatives and mediates delivery of 5-methyltetrahydrofolate to the interior of cells.
 -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By similarity).
 -1- PTM: Eight disulfide bonds are present (Probable).
 -1- SIMILARITY: Belongs to the folate receptor family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 CC EMBL: M64817; AAA37599.1; -
 DR EMBL: L25338; AAA37594.1; -
 DR PIR: B40969; B40969.
 DR MGI: 95569; Folr2.
 DR InterPro: IPR004269; Folt-chemrecept.
 DR Pfam: PR03024; Folate_rec1.
 KW Receptor; Glycoprotein; Signal; Placenta; Folate-binding; Membrane; GPI-anchor; Multigene family; Lipoprotein.
 FT SIGNAL 1 20
 FT CHAIN 21 227 FOLATE RECEPTOR BETA.
 FT PROPEP 228 251 REMOVED IN MATURE FORM (POTENTIAL).
 FT LIPID 227 227 GPI-anchor amidated serine (POTENTIAL).
 FT CARBOHYD 62 62 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 193 193 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 251 AA; 28821 MW; 8404EACB1BFEC7 CRC64;

Query Match 41.4%; Score 41; DB 1; Length 251;
 Best Local Similarity 75.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 8 CHQIEYYP 15
 Db 167 CHTFEYYP 174

RESULT 15
 ID IFT4 HUMAN STANDARD; PRT; 490 AA.
 AC 014879; 099634; 09BSK7;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Interferon-induced protein with tetratricopeptide repeats 4 (IFIT-4)
 DE (Interferon-induced 60 kDa protein) (IFI-60K) (ISG-60) (CIG49)
 DE (Relic acid-induced gene G protein) (RIG-G).
 GN IFT4 OR IFI60.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RX MEDLINE=99047533; PubMed=9391139;
 RA Zhu H., Cong J.P., Shenk T.;
 RT "Use of differential display analysis to assess the effect of human RT cytomagalovirus infection on the accumulation of cellular RNAs: induction of interferon-responsive RNAs.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:13985-13990(1997).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99047533; PubMed=9828129;
 RA de Veer M.J., Sim H., Whistock J.C., Devenish R.J., Ralph S.J.;
 RT "IFI60/ISG60/IFT4, a new member of the human IFI54/IFT2 family of RT interferon-stimulated genes.";
 RL Genomics 54:267-277(1998).
 [3]
 RP SEQUENCE FROM N.A.
 RA Yu M., Tong J., Mao M., Chen S., Chen Z.;
 RT "RIG-G, a novel gene induced by ATRA in acute promyelocytic leukemia cells, is a new member of the ISG family.";
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant P., Prange C., Brownstein M.J., Ussid T.B., Toshiyuki S., Carninci P., Mallory S.J., Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.W., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Alakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E., Jones S.J.M., Maita W.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 -1- SIMILARITY: Belongs to the IFIT family.
 -1- SIMILARITY: Contains 8 TPR repeats.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 CC EMBL: AF026939; AAB95160.1; -
 DR EMBL: AF083470; AAC63524.1; -
 DR EMBL: U52513; AAB40606.1; -
 DR EMBL: BC001383; AAB01383.1; -
 DR EMBL: BC004977; AAH04977.1; -
 DR Genew; HGNC:5411; IFT4.
 DR MIM: 604650; -
 DR InterPro: IPR008941; TPR-like.

```

DR Interpro; IPR001440; TPR.
DR Pfam; PF00515; TPR; 5.
DR SMART; SM00028; TPR; 4.
KW Repeat; TPR repeat; Interferon induction.
FT REPEAT 51 84 TPR 1.
FT REPEAT 94 127 TPR 2.
FT REPEAT 136 169 TPR 3.
FT REPEAT 172 206 TPR 4.
FT REPEAT 207 240 TPR 5.
FT REPEAT 241 274 TPR 6.
FT REPEAT 415 448 TPR 7.
FT REPEAT 450 481 TPR 8.
FT CONFLICT 44 44 F -> S (IN REF. 4; AAH04977).
FT CONFLICT 359 359 Q -> QQ (IN REF. 2).
FT CONFLICT 435 435 MISSING (IN REF. 2).
SQ SEQUENCE 490 AA: 55984 MW; B9F042D4DF7151D2 CRC64;

Query Match 41.4%; Score 41; DB 1; Length 490;
Best local Similarity 43.8%; Pred. No. 26;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 3 LEAKICHOIEYFQDF 18
Db 34 LEDRVCNQIEFLNTEF 49

```

Search completed: September 10, 2004, 17:53:06
 Job time : 8.24022 secs

RESULT 14

US-10-054-611-26
Sequence 26, Application US/10054611
Publication No. US20030059787A1
GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin
Andrews, William H.

TITLE OF INVENTION: No. US20030059787A1e1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,611
FILING DATE: 18-Jan-2002
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/854,050
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20030059787A1 Relevant
TOPOLOGY: No. US20030059787A1 Relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 26:

US-10-054-611-26

Query Match 67.9%; Score 57; DB 14; Length 39;
Best Local Similarity 90.9%; Pred. No. 0.021;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 IIRQVEYFGD 15
DB 1 IIRQVEYFGD 11

RESULT 15

US-10-170-385-477
Sequence 477, Application US/10170385
Publication No. US20030203372A1
GENERAL INFORMATION:

APPLICANT: Ward, Neil Raymond
Mundy, Christopher Robert

APPLICANT: Kan, On
Harris, Robert Alan
White, Jonathan
Binley, Katie Mary
Rayner, William Nigel
Naylor, Stuart
Kingsman, Susan Mary
Krige, David

TITLE OF INVENTION: ANALYSIS METHOD
FILE REFERENCE: 53268200100
CURRENT APPLICATION NUMBER: US/10/170,385
CURRENT FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: PCT/GB02/01662
PRIOR FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: PCT/GB01/05458
PRIOR FILING DATE: 2001-12-10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 477
LENGTH: 408
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-170-385-477

Query Match 67.9%; Score 57; DB 12; Length 408;
Best Local Similarity 66.7%; Pred. No. 0.25;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ERRIQVEYFGDF 16
DB 14 EAKICHQIEYFGDF 28

Search completed: September 10, 2004, 18:11:58
Job time : 36.9436 secs

STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 216:
US-10-325-810-216

Query Match 67.9%; Score 57; DB 12; Length 39;
Best Local Similarity 90.9%; Pred. No. 0.021;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 IIRQVEYFGD 15
|:|||||
Db 1 IIRQVEYFGD 11

RESULT 12
US-10-053-758-26
; Sequence 26, Application US/10053758
; Publication No. US20030032075A1
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
TITLE OF INVENTION: No. US20030032075A1el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/053,758
FILING DATE: 18-Jan-2002
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20030032075A1 Relevant
TOPOLOGY: No. US20030032075A1 Relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-10-053-758-26

Query Match 67.9%; Score 57; DB 14; Length 39;
Best Local Similarity 90.9%; Pred. No. 0.021;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 IIRQVEYFGD 15
|:|||||
Db 1 IIRQVEYFGD 11

RESULT 13
US-10-054-295-26
; Sequence 26, Application US/10054295
; Publication No. US20030044953A1
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
TITLE OF INVENTION: No. US20030044953A1el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,295
FILING DATE: 18-Jan-2002
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/854,050
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20030044953A1 Relevant
TOPOLOGY: No. US20030044953A1 Relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-10-054-295-26

Query Match 67.9%; Score 57; DB 14; Length 39;
Best Local Similarity 90.9%; Pred. No. 0.021;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 IIRQVEYFGD 15
|:|||||
Db 1 IIRQVEYFGD 11

MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-766-253-26

Query Match
Best Local Similarity 67.9%; Score 57; DB 9; Length 39;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 IIRQVEYIFGD 15
DB 1 IIRQVEYIFGD 11

RESULT 10
US-09-438-486-26
Sequence 26, Application US/09438486
Publication No. US2003009019A1
GENERAL INFORMATION:
APPLICANT: Cecch, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin H.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. US2003009019A1el Telomerase
NUMBER OF SEQUENCES: 223
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/438,486
FILING DATE: 12-NOV-1999
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002931US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide

US-09-438-486-26

Query Match
Best Local Similarity 67.9%; Score 57; DB 10; Length 39;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 IIRQVEYIFGD 15
DB 1 IIRQVEYIFGD 11

RESULT 11
US-10-325-810-216
Sequence 216, Application US/10325810
Publication No. US20030204069A1
GENERAL INFORMATION:
APPLICANT: Cecch, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin H.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 633
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/325,810
FILING DATE: 20-Dec-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/402,181
FILING DATE: 29-Sep-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ausehus, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 216:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid

RESULT 7
US-09-836-073-14
Sequence 14, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Daegupta, Asim
APPLICANT: Das, S.
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 22002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 18
TYPE: PRT
ORGANISM: Bovine
US-09-836-073-14

Query Match 67.9%; Score 57; DB 9; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.0095;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ERAIIRQVEYFGDF 16
Db 4 EAKICHQIEYFGDF 18

RESULT 8
US-09-843-676-26
Sequence 26, Application US/09843676
Patent No. US20020164786A1
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin
Andrews, William H.
TITLE OF INVENTION: No. US20020164786A1e1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/843,676
FILING DATE: 26-Apr-2001
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20020164786A1 Relevant
MOLECULE TYPE: peptide
TOPOLOGY: No. US20020164786A1 Relevant
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-843-676-26

Query Match 67.9%; Score 57; DB 9; Length 39;
Best Local Similarity 90.9%; Pred. No. 0.021;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 IIRQVEYFGD 15
Db 1 IIRQVEYFGD 11

RESULT 9
US-09-766-253-26
Sequence 26, Application US/09766253
Publication No. US20020187471A1
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin
Andrews, William H.
TITLE OF INVENTION: No. US20020187471A1e1 Telomerase
NUMBER OF SEQUENCES: 171
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/766,253
FILING DATE: 19-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/846,017
FILING DATE: 1997-04-25
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002920US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant

```

; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-4

Query Match      73.8%; Score 62; DB 9; Length 18;
Best Local Similarity 68.8%; Pred. No. 0.0013;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1 0ERAIRQVEYFFGDF 16
DB      3 0EAKICHQLEYYFGDF 18

RESULT 3
US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match      69.0%; Score 58; DB 9; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.0064;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      2 0ERAIRQVEYFFGDF 16
DB      4 0EAIICQIIEYFFGDF 18

RESULT 4
US-09-836-073-13
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 17
```

```

; TYPE: PRT
; ORGANISM: Mouse
US-09-836-073-13

Query Match      67.9%; Score 57; DB 9; Length 17;
Best Local Similarity 66.7%; Pred. No. 0.0089;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      2 0ERAIRQVEYFFGDF 16
DB      3 0EAKICHQLEYYFGDF 17

RESULT 5
US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-1

Query Match      67.9%; Score 57; DB 9; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.0095;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      2 0ERAIRQVEYFFGDF 16
DB      4 0EAKICHQLEYYFGDF 18

RESULT 6
US-09-836-073-5
; Sequence 5, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-5

Query Match      67.9%; Score 57; DB 9; Length 18;
Best Local Similarity 62.5%; Pred. No. 0.0095;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      1 0ERAIRQVEYFFGDF 16
DB      3 0EAKICHQLEYYFGDF 18
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:52:06 : Search time 35.8436 Seconds
(without alignments)
143.151 Million cell updates/sec

Title: US-09-836-073-19

Sequence: 1 QERRAIIRQVEYFGDF 16

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
- 4: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
- 5: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
- 6: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
- 7: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84	100.0	16	US-09-836-073-19	Sequence 19, Appl
2	62	73.8	18	US-09-836-073-4	Sequence 4, Appl
3	58	69.0	18	US-09-836-073-2	Sequence 2, Appl
4	57	67.9	17	US-09-836-073-13	Sequence 13, Appl
5	57	67.9	18	US-09-836-073-1	Sequence 1, Appl
6	57	67.9	18	US-09-836-073-5	Sequence 5, Appl
7	57	67.9	18	US-09-836-073-14	Sequence 14, Appl
8	57	67.9	18	US-09-836-073-26	Sequence 26, Appl
9	57	67.9	39	US-09-766-253-26	Sequence 26, Appl
10	57	67.9	39	US-09-438-486-26	Sequence 26, Appl
11	57	67.9	39	US-10-325-810-216	Sequence 216, Appl
12	57	67.9	39	US-10-053-758-26	Sequence 26, Appl
13	57	67.9	39	US-10-054-295-26	Sequence 26, Appl
14	57	67.9	39	US-10-054-611-26	Sequence 26, Appl
15	57	67.9	408	US-10-170-385-477	Sequence 477, App

16	57	67.9	460	12	US-09-925-298-695	Sequence 695, App
17	57	67.9	460	14	US-10-102-806-695	Sequence 695, App
18	57	67.9	460	15	US-10-264-049-2643	Sequence 2643, App
19	55	65.5	488	12	US-10-424-559-272690	Sequence 272690, App
20	54	64.3	18	9	US-09-836-073-15	Sequence 15, Appl
21	54	64.3	21	15	US-10-376-121A-20	Sequence 20, Appl
22	53	63.1	411	14	US-10-177-478-8	Sequence 8, Appl
23	51	60.7	18	9	US-09-836-073-9	Sequence 9, Appl
24	50	59.5	18	9	US-09-836-073-17	Sequence 17, Appl
25	50	59.5	38	12	US-10-325-810-214	Sequence 214, App
26	50	59.5	143	12	US-10-424-599-254661	Sequence 254661, App
27	50	59.5	395	12	US-10-424-599-254664	Sequence 254664, App
28	49	58.3	18	9	US-09-836-073-11	Sequence 11, Appl
29	49	58.3	18	9	US-09-836-073-12	Sequence 12, Appl
30	49	58.3	376	16	US-10-767-701-45524	Sequence 45524, A
31	49	58.3	405	16	US-10-437-963-134637	Sequence 134637, A
32	49	58.3	467	16	US-10-437-963-155751	Sequence 155751, A
33	48	57.1	18	9	US-09-836-073-10	Sequence 10, Appl
34	48	57.1	38	9	US-09-843-676-25	Sequence 25, Appl
35	48	57.1	38	9	US-09-766-253-25	Sequence 25, Appl
36	48	57.1	38	10	US-09-438-486-25	Sequence 25, Appl
37	48	57.1	38	12	US-10-325-810-215	Sequence 215, Appl
38	48	57.1	38	14	US-10-053-758-25	Sequence 25, Appl
39	48	57.1	38	14	US-10-054-295-25	Sequence 25, Appl
40	48	57.1	38	14	US-10-054-611-25	Sequence 25, Appl
41	48	57.1	148	16	US-10-767-701-53447	Sequence 53447, A
42	48	57.1	420	16	US-10-437-963-128072	Sequence 128072, A
43	48	57.1	453	16	US-10-437-963-179489	Sequence 179489, A
44	47	56.0	48	12	US-10-424-599-269667	Sequence 269667, A
45	47	56.0	192	16	US-10-767-701-40959	Sequence 40959, A

ALIGNMENTS

RESULT 1
US-09-836-073-19
Sequence 19, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Dasgupta, Asim
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 22002054822
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: US/09/836,073
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 19
LENGTH: 16
TYPE: PRT
ORGANISM: Drosophila
US-09-836-073-19

Query Match 100.0%; Score 84; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QERRAIIRQVEYFGDF 16
Db 1 QERRAIIRQVEYFGDF 16

RESULT 2
US-09-836-073-4
Sequence 4, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Dasgupta, Asim

FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-215

Query Match 57.1%; Score 48; DB 3; Length 38;
Best Local Similarity 72.7%; Pred. No. 0.082;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IIRQVEYYFGD 15
| : : : : :
Db 1 ICEQIEYYFGD 11

RESULT 15
US-08-854-050-25
Sequence 25, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin H.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6261836el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-25

Query Match 57.1%; Score 48; DB 3; Length 38;
Best Local Similarity 72.7%; Pred. No. 0.082;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IIRQVEYYFGD 15
| : : : : :
Db 1 ICEQIEYYFGD 11

Search completed: September 10, 2004, 18:05:12
Job time: 14.1229 secs

FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 214:
US-09-721-456-214

Query Match 59.5%; Score 50; DB 4; Length 38;
Best Local Similarity 75.0%; Pred. NO. 0.036;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 IIRQVEYFGDF 16
DB 1 ICHQXEYFGDF 12

RESULT 13
US-08-851-843A-25
Sequence 25, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-25

Query Match 57.1%; Score 48; DB 3; Length 38;
Best Local Similarity 72.7%; Pred. NO. 0.082;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IIRQVEYFGD 15
DB 1 ICEQIEYFGD 11

RESULT 14
US-08-974-549A-215
Sequence 215, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017

APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/117618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/117885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-214

Query Match
Best Local Similarity 59.5%; Score 50; DB 3; Length 38;
Best Local Similarity 75.0%; Pred. No. 0.036;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 IIRQVEYYFGDF 16
DB 1 ICHQXVEYYFGDF 12

RESULT 11
US-09-402-181B-214
Sequence 214, Application US/09402181B
Patent No. 6610839
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 633
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,181B
FILING DATE: 29-Sep-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/117885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ausubius, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 214:
US-09-402-181B-214

Query Match
Best Local Similarity 59.5%; Score 50; DB 4; Length 38;
Best Local Similarity 75.0%; Pred. No. 0.036;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 IIRQVEYYFGDF 16
DB 1 ICHQXVEYYFGDF 12

RESULT 12
US-09-721-456-214
Sequence 214, Application US/09721456
Patent No. 6617110
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/721,456
FILING DATE: 22-NOV-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017

FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 216:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 216:
US-09-721-456-216

Query Match 67.9%; Score 57; DB 4; Length 39;
Best Local Similarity 90.9%; Pred. No. 0.0021;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 IIRQVEYFGD 15
|:|||||
Db 1 ILRQVEYFGD 11

RESULT 9
US-08-475-955-20
Sequence 20, Application US/08475955
Patent No. 6641813
GENERAL INFORMATION:
APPLICANT: Harley, John
TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS OF
NUMBER OF SEQUENCES: 218
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center, 1250 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,955
FILING DATE: June 7, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/867,819
FILING DATE: April 13, 1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/648,205
FILING DATE: January 31, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/472,947
FILING DATE: January 31, 1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRF114CIP(2)DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-873-8795

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Binding-site
LOCATION: 8..15
US-08-475-955-20

Query Match 64.3%; Score 54; DB 4; Length 21;
Best Local Similarity 75.0%; Pred. No. 0.0035;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IIRQVEYFGDF 16
|:|||||
Db 1 ICHQIEYFGDF 12

RESULT 10
US-08-974-549A-214
Sequence 214, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cecch, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morlin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:

REFERENCE/DOCKET NUMBER: 015389-002930US
REGISTRATION NUMBER: 42,271
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-430-323-26

Query Match 67.9%; Score 57; DB 4; Length 39;
Best Local Similarity 90.9%; Pred. No. 0.0021;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 IIRQVEYRFGD 15
DB 1 IIRQVEYRFGD 11

RESULT 7
US-09-402-181B-216
Sequence 216, Application US/09402181B
Patent No. 6610839
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 633
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,181B
FILING DATE: 29-Sep-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:

NAME: Ausehus, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 216:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 216:
US-09-402-181B-216

Query Match 67.9%; Score 57; DB 4; Length 39;
Best Local Similarity 90.9%; Pred. No. 0.0021;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 IIRQVEYRFGD 15
DB 1 IIRQVEYRFGD 11

RESULT 8
US-09-721-456-216
Sequence 216, Application US/09721456
Patent No. 6617110
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/721,456
FILING DATE: 22-No. 6617110-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA: WO PCT/US97/17685
APPLICATION NUMBER: WO PCT/US97/17685
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 216:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-216

Query Match 67.9%; Score 57; DB 3; Length 39;
Best Local Similarity 90.9%; Pred. No. 0.0021;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 IIRQVEYFGD 15
Db 1 IIRQVEYFGD 11

RESULT 5
US-08-854-050-26
Sequence 26, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6261836el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-26

Query Match 67.9%; Score 57; DB 3; Length 39;
Best Local Similarity 90.9%; Pred. No. 0.0021;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 IIRQVEYFGD 15
Db 1 IIRQVEYFGD 11

RESULT 6
US-09-430-323-26
Sequence 26, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-OCT-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429

PRIOR FILING DATE: 1997-10-06
PRIOR APPLICATION NUMBER: 08/321,427
PRIOR FILING DATE: 1994-10-11
PRIOR APPLICATION NUMBER: 60/086,527
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 18
TYPE: PR1
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: B-LAP
US-09-316-630-4

Query Match 67.9%; Score 57; DB 3; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.00085;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 EBAIRQVEYFGDF 16
DB 4 EAKICHQIEYFGDF 18

RESULT 3
US-08-851-843A-26
Sequence 26, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-26

Query Match 67.9%; Score 57; DB 3; Length 39;
Best Local Similarity 90.9%; Pred. No. 0.0021;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 IIRQVEYFGD 15
DB 1 IIRQVEYFGD 11

RESULT 4
US-08-974-549A-216
Sequence 216, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:43:01 ; Search time 14.1229 Seconds
(without alignments)
58,488 Million cell updates/sec

Title: US-09-836-073-19

Perfect score: 84
Sequence: 1,GERAIIRROYEYFGDF 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/1aa/PTUS_COMB.pep.*
6: /cgn2_6/prodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	67.9	18	3 US-09-316-630-3	Sequence 3, Appl1
2	57	67.9	18	3 US-09-316-630-4	Sequence 4, Appl1
3	57	67.9	39	3 US-08-851-843A-26	Sequence 26, Appl1
4	57	67.9	39	3 US-08-974-549A-216	Sequence 26, Appl1
5	57	67.9	39	3 US-08-854-050-26	Sequence 26, Appl1
6	57	67.9	39	4 US-09-430-323-26	Sequence 26, Appl1
7	57	67.9	39	4 US-09-402-181B-216	Sequence 216, Appl1
8	57	67.9	39	4 US-09-721-456-216	Sequence 216, Appl1
9	54	64.3	21	4 US-08-475-955-20	Sequence 20, Appl1
10	50	59.5	38	4 US-08-974-549A-214	Sequence 214, Appl1
11	50	59.5	38	4 US-09-402-181B-214	Sequence 214, Appl1
12	50	59.5	38	4 US-09-721-456-214	Sequence 214, Appl1
13	48	57.1	38	4 US-08-851-843A-25	Sequence 25, Appl1
14	48	57.1	38	3 US-08-974-549A-215	Sequence 25, Appl1
15	48	57.1	38	3 US-08-854-050-25	Sequence 25, Appl1
16	48	57.1	38	4 US-09-430-323-25	Sequence 25, Appl1
17	48	57.1	38	4 US-09-402-181B-215	Sequence 215, Appl1
18	48	57.1	38	4 US-09-721-456-215	Sequence 215, Appl1
19	43	51.2	37	3 US-08-851-843A-24	Sequence 24, Appl1
20	43	51.2	37	3 US-08-854-050-24	Sequence 24, Appl1
21	43	51.2	37	4 US-09-430-323-24	Sequence 24, Appl1
22	42	50.0	38	3 US-08-851-843A-27	Sequence 27, Appl1
23	42	50.0	38	3 US-08-974-549A-217	Sequence 27, Appl1
24	42	50.0	38	3 US-08-854-050-27	Sequence 27, Appl1
25	42	50.0	38	4 US-09-430-323-27	Sequence 27, Appl1
26	42	50.0	38	4 US-09-402-181B-217	Sequence 217, Appl1
27	42	50.0	38	4 US-09-721-456-217	Sequence 217, Appl1

28	42	50.0	167	4	US-09-134-000C-3435	Sequence 3435, Appl1
29	41	48.8	74	4	US-09-439-554-6	Sequence 6, Appl1
30	41	48.8	141	4	US-09-540-236-2332	Sequence 2332, Appl1
31	41	48.8	292	4	US-09-439-554-24	Sequence 24, Appl1
32	40.5	48.2	2710	2	US-08-568-459A-12	Sequence 12, Appl1
33	40.5	48.2	2710	2	US-08-487-826B-12	Sequence 12, Appl1
34	40.5	48.2	2710	4	US-09-210-288-12	Sequence 12, Appl1
35	40.5	48.2	3060	2	US-08-487-826B-14	Sequence 14, Appl1
36	40.5	48.2	3060	2	US-08-487-826B-12	Sequence 12, Appl1
37	40	47.6	376	4	US-09-489-039A-11743	Sequence 11743, Appl1
38	40	47.6	506	4	US-09-672-785-2	Sequence 2, Appl1
39	40	47.6	512	4	US-09-672-785-8	Sequence 8, Appl1
40	40	47.6	1076	4	US-09-470-443-6	Sequence 6, Appl1
41	40	47.6	1145	4	US-09-470-443-2	Sequence 2, Appl1
42	40	47.6	1145	4	US-09-470-443-4	Sequence 4, Appl1
43	40	47.6	3571	4	US-09-911-842A-2	Sequence 2, Appl1
44	39	46.4	69	4	US-09-621-976-7385	Sequence 7385, Appl1
45	39	46.4	319	4	US-09-252-991A-29238	Sequence 29238, Appl1

ALIGNMENTS

```

RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; TITLE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: IAP
US-09-316-630-3
Query Match 67.9%; Score 57; DB 3; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.00085;
Mat:hes 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Cy 2 ERAIIRROYEYFGDF 16
Db 4 ERAIIRROYEYFGDF 18
RESULT 2
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; TITLE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953

```


modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 408 AA:

Query Match 67.9%; Score 57; DB 7; Length 408;
Best Local Similarity 66.7%; Pred. No. 0.18;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ERAIRQVEYFGDF 16
| : |||||
Db 14 EAKICHQIEYFGDF 28

RESULT 14
ADD46272
ID ADD46272 standard; protein; 408 AA.

AC ADD46272;

DT 29-JAN-2004 (first entry)

DE Human Protein P05455, SEQ ID NO 11947.

XX Human; pain; neuronal tissue; gene therapy;

KM spinal segmental nerve injury; chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

OS Homo sapiens.

XX WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-033347P.

PA (GEHO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

DR GENBANK; P05455.

XX Claim 1; Page; 1017p; English.

The invention discloses a composition comprising two or more isolated rat CC or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also CC claimed are a vector comprising the novel polynucleotide, a host cell CC comprising the vector, a method for identifying a nucleotide sequence CC which is differentially regulated in an animal subjected to pain and a CC kit to perform the method, an array, a method for identifying an agent CC that increases or decreases the expression of the polynucleotide sequence CC that is differentially expressed in neuronal tissue of a first animal CC subjected to pain, a method for identifying a compound which regulates CC the expression of a polynucleotide sequence which is differentially CC expressed in an animal subjected to pain, a method for identifying a CC compound that regulates the activity of one or more of the CC polynucleotides, a method for producing a pharmaceutical composition, a

method for identifying a compound or small molecule that regulates the CC activity in an animal of one or more of the polypeptides given in the CC specification, a method for identifying a compound useful in treating CC pain and a pharmaceutical composition comprising the one or more CC polypeptides or their antibodies. The polynucleotide or the compound that CC modulates its activity is useful for preparing a medicament for treating CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene CC therapy). The sequence presented is a human protein (shown in Table 2 of CC the specification) which is differentially expressed during pain. Note: CC the sequence data for this patent did not form part of the printed CC specification, but was obtained in electronic form directly from WIPO at CC ftp.wipo.int/pub/published_pct_sequences.

Sequence 408 AA:

Query Match 67.9%; Score 57; DB 7; Length 408;
Best Local Similarity 66.7%; Pred. No. 0.18;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ERAIRQVEYFGDF 16
| : |||||
Db 14 EAKICHQIEYFGDF 28

RESULT 15
ADE63995
ID ADE63995 standard; protein; 408 AA.

AC ADE63995;

DT 29-JAN-2004 (first entry)

DE Human Protein P05455, SEQ ID NO 9941.

XX Human; pain; neuronal tissue; gene therapy;

KM spinal segmental nerve injury; chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

OS Homo sapiens.

XX WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-033347P.

PA (GEHO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

DR GENBANK; P05455.

XX Claim 1; Page; 1017p; English.

The invention discloses a composition comprising two or more isolated rat CC or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also CC claimed are a vector comprising the novel polynucleotide, a host cell CC comprising the vector, a method for identifying a nucleotide sequence CC which is differentially regulated in an animal subjected to pain and a CC kit to perform the method, an array, a method for identifying an agent CC that increases or decreases the expression of the polynucleotide sequence CC that is differentially expressed in neuronal tissue of a first animal

CC autoimmune patients. La and Ro antigens sometimes reside on the same
 CC cellular ribonucleoprotein particle, most La patients contain some Ro
 CC antibodies and vice versa. La cDNA has been isolated from a human liver
 CC library. (Updated on 25-MAR-2003 to correct PF field.)
 XX

SQL Sequence 408 AA;

Query Match 67.9%; Score 57; DB 2; Length 408;
 Best Local Similarity 66.7%; Pred. No. 0.18;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 ERAIRQVEYFYRGDF 16
 | | : | | | | | | | |
 Db 14 EAKICHQIEYFYRGDF 28

RESULT 12

ABP65252
 ID ABP65252 standard; protein; 408 AA.

AC ABP65252;

DT 12-NOV-2002 (first entry)

DE Hypoxia-regulated protein #126.

XX Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;
 KW antiinflammatory; vulnery; gynecological; ophthalmological; vaccine;
 KW hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;
 KW ischaemic condition; reperfusion injury; retinopathy; neonatal stress;
 KW preclapmsia; atherosclerosis; inflammatory condition; wound healing;
 KW inflammation; erythropoiesis; hair loss; human.
 XX

OS Homo sapiens.

PN WO200246465-A2.

PD 13-JUN-2002.

PF 10-DEC-2001; 2001WO-GB005458.

XX 08-DEC-2000; 2000GB-00030076.

PR 08-FEB-2001; 2001GB-00003156.

PR 25-OCT-2001; 2001GB-00025666.

(OXFO-) OXFORD BIOMEDICA UK LTD.

PI White J, Mundy CR, Ward NR, Krige D, Kingsman SM, Harris RA;

PI Rayner WN;

DR WPI; 2002-627238/67.

PT Identifying a gene involved in disease for treating hypoxia-regulated
 PT conditions, comprises comparing the transcriptome/proteome of two cell
 PT types under different conditions and identifying a differentially
 PT regulated gene.
 XX

PS Claim 35; Page 425; 538pp; English.

CC The present invention relates to methods for identifying genes and
 CC proteins that are implicated in a specific disease or physiological
 CC condition. The method comprises comparing the transcriptome/proteome of a
 CC specialised cell type implicated in a disease or condition with that of a
 CC second specialised cell type, under two experimental conditions, and
 CC identifying a gene that is differentially regulated in the two
 CC specialised cell types under experimental conditions. ABV77873-ABV78116
 CC and ABP65061-ABP65257 were identified using the methods of the invention.
 CC The coding sequences and proteins are useful for treating a disease in a
 CC patient, for manufacture of a medicament for treating hypoxia-regulated
 CC conditions, and for regulating tumorigenesis, angiogenesis, apoptosis,
 CC biological response to hypoxia conditions, or hypoxic-associated
 CC pathology in a patient. The coding sequences and proteins are also useful
 CC for monitoring the therapeutic treatment of a disease or physiological

CC condition, such as cancer, ischaemic conditions, reperfusion injury,
 CC retinopathy, neonatal stress, preclapmsia, atherosclerosis, inflammatory
 CC conditions, wound healing, inflammation, erythropoiesis or hair loss
 XX

SQL Sequence 408 AA;

Query Match 67.9%; Score 57; DB 5; Length 408;
 Best Local Similarity 66.7%; Pred. No. 0.18;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 ERAIRQVEYFYRGDF 16
 | | : | | | | | | | |
 Db 14 EAKICHQIEYFYRGDF 28

RESULT 13

ADE63991

ID ADE63991 standard; protein; 408 AA.

AC ADE63991;

DT 29-JAN-2004 (first entry)

DE Human Protein P05455; SEQ ID NO 9937.

XX Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SN1; Chung.
 XX

OS Homo sapiens.

PN WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-033347P.

(GEHO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

DR WPI; 2003-268312/26.

DR GENBANK; P05455.

PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX

PS Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that

KW viral life cycle; antiviral therapy; pharmaceutical; veterinary;
 KW agricultural; horticultural; virucide; bovine.
 XX Bovinae.
 OS
 XX WO200283858-A2.
 PN
 XX 24-OCT-2002.
 PD
 XX
 XX 12-APR-2002; 2002WO-US011589.
 PF
 XX 16-APR-2001; 2001US-00836073.
 PR
 XX (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.
 PA
 XX Dasgupta A, Das S, Baidya N;
 PI WPI; 2003-058634/05.
 DR
 XX
 XX New compound containing acidic and aromatic amino acids, useful as
 PT antiviral therapy in pharmaceutical, veterinary or
 PT agricultural/horticultural applications.
 PT
 XX Disclosure; Page 6; 19pp; English.
 XX
 CC The present invention relates to peptides and methods of inhibiting the
 CC replication of viruses that utilise internal ribosome entry site (IRES)
 CC initiated translation, and/or inhibiting viruses that utilise the Ia
 CC antigen protein (IAP) in any phase of their life cycle. The peptides of
 CC the invention compete with IAP and inhibit the utilisation of various
 CC biochemical and physiological functions of IAP required for a productive
 CC life cycle. The methods and compositions are useful as antiviral therapy
 CC in pharmaceutical, veterinary or agricultural/horticultural applications.
 CC ABG72101-ABG72119 represent peptides useful as antiviral agents
 CC
 XX
 SQ Sequence 18 AA;
 Query Match 67.9%; Score 57; DB 6; Length 18;
 Best Local Similarity 66.7%; Pred. NO. 0.0048; 4; Indels 0; Gaps 0;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 2 ERAIIRQVEYFGDF 16
 Db 4 EAKICHQIEYFGDF 18
 DE
 XX
 XX RESULT 10
 XX AAG01351
 ID AAG01351 standard; protein; 92 AA.
 AC
 AC AAG01351;
 DT
 XX 06-OCT-2000 (first entry)
 DT
 XX
 DE Human secreted protein, SEQ ID NO: 5432.
 XX
 KW Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 OS
 XX EPI033401-A2.
 PN
 XX 06-SEP-2000.
 PD
 XX 21-FEB-2000; 2000EP-00200610.
 PF
 XX 26-FEB-1999; 99US-0122487P.
 PR
 XX
 XX (GENSET) GENSET.
 PA
 XX Dunas Milne Edwards J, Duclert A, Giordano J;
 PI
 XX

DR WPI; 2000-500381/45.
 DR N-PSDB; AAC01357.
 XX
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
 PT
 XX Claim 13; SEQ ID NO.5432; 71pp + Sequence Listing; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number of
 CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
 CC prepared from total human RNAs or polyA+ RNAs derived from 30 different
 CC tissues. EST sequences usually correspond mainly to the 3' untranslated
 CC region (UTR) of the mRNA because they are often obtained from oligo-dT
 CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
 CC sequences derived from the 5' ends of mRNAs and even in those cases where
 CC longer cDNA sequences have been obtained, the full 5' UTR is rarely
 CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can
 CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
 CC are also used in diagnostic, forensic, gene therapy and chromosome
 CC mapping procedures. They are used to obtain upstream regulatory sequences
 CC and to design expression and secretion vectors
 CC
 XX
 SQ Sequence 92 AA;
 Query Match 67.9%; Score 57; DB 3; Length 92;
 Best Local Similarity 66.7%; Pred. NO. 0.032;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 2 ERAIIRQVEYFGDF 16
 Db 14 EAKICHQIEYFGDF 28
 DE
 XX
 XX RESULT 11
 XX AAW03716
 ID AAW03716 standard; protein; 408 AA.
 AC
 AC AAW03716;
 DT
 XX 25-MAR-2003 (revised)
 DT
 XX 12-MAR-1997 (first entry)
 DT
 XX
 DE Human autoantigen Ia(SS-B).
 XX
 KW Autoimmune disease; Ia autoantigen; Sjogren's syndrome;
 KW systemic lupus erythematosus; diagnosis.
 KW
 OS Homo sapiens.
 OS
 XX US5541291-A.
 PN
 XX 30-JUL-1996.
 PD
 XX 27-MAY-1987; 87US-00054871.
 PF
 XX 31-DEC-1984; 84US-00687908.
 PR
 XX (UYDU-) UNIV DUKE.
 PA
 XX
 XX Keene JD;
 PI
 XX WPI; 1996-362015/36.
 DR
 XX Auto-antigen U2-RNP, associated with severe polymyositis-scleroderma
 PT overlap syndrome - useful for diagnosis and treatment of autoimmune
 PT diseases.
 PT
 XX Disclosure; Col 15-16; 21pp; English.
 XX
 CC The human lupus antigen (Ia) is diagnostic for Sjogren's syndrome, as
 CC well as occurring in systemic lupus erythematosus patients. The Ia protein
 CC is clinically related to the Ro protein that is highly common among

CC coxsackie virus, encephalomyocarditis virus, foot-and-mouth disease
 CC virus, echo virus, hepatitis C virus, infectious bronchitis virus, duck
 CC and human hepatitis B virus, and vesicular stomatitis virus. The peptide
 CC also inhibits replication of the above viruses. The LAP peptide
 CC selectively inhibits viral protein translation, and is therefore not
 CC toxic to the host cell
 XX

SQ Sequence 18 AA;

Query Match 67.9%; Score 57; DB 3; Length 18;
 Best Local Similarity 66.7%; Pred. No. 0.0048;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ERAIRIQVEYFGDF 16
 DB 4 EAKICHOIEYFGDF 18

RESULT 7

ID ABG72101 standard; peptide; 18 AA.

AC ABG72101;

DT 28-JAN-2003 (first entry)

DE Viral replication inhibiting peptide, LAP.

XX Viral replication inhibitor; IRES initiated translation; LAP;

KM Internal ribosome entry site initiated translation; La antigen protein;

KM viral life cycle; antiviral therapy; pharmaceutical; veterinary;

KM agricultural; horticultural; virucide.

OS Unidentified.

PN WO200283858-A2.

PD 24-OCT-2002.

PF 12-APR-2002; 2002WO-US011589.

PR 16-APR-2001; 2001US-00836073.

RA (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.

PI Dasgupta A, Das S, Baidya N;

DR WPI; 2003-058634/05.

XX New compound containing acidic and aromatic amino acids, useful as
 PT antiviral therapy in pharmaceutical, veterinary or
 PT agricultural/horticultural applications.

PS Claim 1; Page 15; 19pp; English.

CC The present invention relates to peptides and methods of inhibiting the
 CC replication of viruses that utilise internal ribosome entry site (IRES)
 CC initiated translation, and/or inhibiting viruses that utilise the La
 CC antigen protein (LAP) in any phase of their life cycle. The peptides of
 CC the invention compete with LAP and inhibit the utilisation of various
 CC biochemical and physiological functions of LAP required for a productive
 CC life cycle. The methods and compositions are useful as antiviral therapy
 CC in pharmaceutical, veterinary or agricultural/horticultural applications.
 CC ABG72101-ABG72119 represent peptides useful as antiviral agents
 XX

SQ Sequence 18 AA;

Query Match 67.9%; Score 57; DB 6; Length 18;
 Best Local Similarity 66.7%; Pred. No. 0.0048;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ERAIRIQVEYFGDF 16
 DB 4 EAKICHOIEYFGDF 18

DB 4 EAKICHOIEYFGDF 18

RESULT 8
 ABG72105
 ID ABG72105 standard; peptide; 18 AA.

AC ABG72105;

DT 28-JAN-2003 (first entry)

DE Viral replication inhibiting peptide, 762.

XX Viral replication inhibitor; IRES initiated translation; LAP;

KM Internal ribosome entry site initiated translation; La antigen protein;

KM viral life cycle; antiviral therapy; pharmaceutical; veterinary;

KM agricultural; horticultural; virucide.

OS Unidentified.

PN WO200283858-A2.

PD 24-OCT-2002.

PF 12-APR-2002; 2002WO-US011589.

PR 16-APR-2001; 2001US-00836073.

RA (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.

PI Dasgupta A, Das S, Baidya N;

DR WPI; 2003-058634/05.

XX New compound containing acidic and aromatic amino acids, useful as
 PT antiviral therapy in pharmaceutical, veterinary or
 PT agricultural/horticultural applications.

PS Example 3; Page 14; 19pp; English.

CC The present invention relates to peptides and methods of inhibiting the
 CC replication of viruses that utilise internal ribosome entry site (IRES)
 CC initiated translation, and/or inhibiting viruses that utilise the La
 CC antigen protein (LAP) in any phase of their life cycle. The peptides of
 CC the invention compete with LAP and inhibit the utilisation of various
 CC biochemical and physiological functions of LAP required for a productive
 CC life cycle. The methods and compositions are useful as antiviral therapy
 CC in pharmaceutical, veterinary or agricultural/horticultural applications.
 CC ABG72101-ABG72119 represent peptides useful as antiviral agents
 XX

SQ Sequence 18 AA;

Query Match 67.9%; Score 57; DB 6; Length 18;
 Best Local Similarity 62.5%; Pred. No. 0.0048;
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 QERAIRIQVEYFGDF 16
 DB 3 QEOXCHOIEYFGDF 18

RESULT 9

ABG72114
 ID ABG72114 standard; peptide; 18 AA.

AC ABG72114;

DT 28-JAN-2003 (first entry)

DE Viral replication inhibiting peptide, BOVINE.

KM Viral replication inhibitor; IRES initiated translation; LAP;

KM Internal ribosome entry site initiated translation; La antigen protein;

PD 24-OCT-2002.
XX
XX
PF 12-APR-2002; 2002WO-US011589.
XX
XX 16-APR-2001; 2001US-00836073.
XX
XX (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.
XX
XX Dasgupta A, Das S, Baidya N;
XX
XX WPI; 2003-058634/05.
XX
XX
XX New compound containing acidic and aromatic amino acids, useful as
XX antiviral therapy in pharmaceutical, veterinary or
XX agricultural/horticultural applications.
XX
XX Claim 10; Page 16; 19pp; English.
XX
XX The present invention relates to peptides and methods of inhibiting the
XX replication of viruses that utilise internal ribosome entry site (IRES)
XX initiated translation, and/or inhibiting viruses that utilise the la
XX antigen protein (LAP) in any phase of their life cycle. The peptides of
XX the invention compete with LAP and inhibit the utilisation of various
XX biochemical and physiological functions of LAP required for a productive
XX life cycle. The methods and compositions are useful as antiviral therapy
XX in pharmaceutical, veterinary or agricultural/horticultural applications.
XX ABG72101-ABG72119 represent peptides useful as antiviral agents
XX
XX Sequence 18 AA;
XX
XX Query Match 69.0%; Score 58; DB 6; Length 18;
XX Best Local Similarity 66.7%; Pred. No. 0.0032;
XX Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
XX
XX QY 2 ERAIIHQVEYFGDF 16
XX | | | | | | | | | |
XX | | | | | | | | | |
XX 4 EAQICQIEYFGDF 18
XX
XX
XX RESULT 5
XX ABG72113
XX ID ABG72113 standard; peptide; 17 AA.
XX
XX AC ABG72113;
XX
XX DT 28-JAN-2003 (first entry)
XX
XX DE Viral replication inhibiting peptide, MOUSE.
XX
XX KM Viral replication inhibitor; IRES initiated translation; LAP;
XX internal ribosome entry site initiated translation; la antigen protein;
XX viral life cycle; antiviral therapy; pharmaceutical; veterinary;
XX agricultural, horticultural; virucide; mouse.
XX
XX OS Mus sp.
XX
XX PN WO200283858-A2.
XX
XX PD 24-OCT-2002.
XX
XX PF 12-APR-2002; 2002WO-US011589.
XX
XX PR 16-APR-2001; 2001US-00836073.
XX
XX PA (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.
XX
XX PI Dasgupta A, Das S, Baidya N;
XX
XX DR WPI; 2003-058634/05.
XX
XX PT New compound containing acidic and aromatic amino acids, useful as
XX antiviral therapy in pharmaceutical, veterinary or
XX agricultural/horticultural applications.

XX
XX Claim 1; Page 15; 19pp; English.
XX
XX The present invention relates to peptides and methods of inhibiting the
XX replication of viruses that utilise internal ribosome entry site (IRES)
XX initiated translation, and/or inhibiting viruses that utilise the la
XX antigen protein (LAP) in any phase of their life cycle. The peptides of
XX the invention compete with LAP and inhibit the utilisation of various
XX biochemical and physiological functions of LAP required for a productive
XX life cycle. The methods and compositions are useful as antiviral therapy
XX in pharmaceutical, veterinary or agricultural/horticultural applications.
XX ABG72101-ABG72119 represent peptides useful as antiviral agents
XX
XX Sequence 17 AA;
XX
XX Query Match 67.9%; Score 57; DB 6; Length 17;
XX Best Local Similarity 66.7%; Pred. No. 0.0045;
XX Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 2 ERAIIHQVEYFGDF 16
XX | | | | | | | | | |
XX | | | | | | | | | |
XX 3 EAKICQIEYFGDF 17
XX
XX
XX RESULT 6
XX AAY52200
XX ID AAY52200 standard; peptide; 18 AA.
XX
XX AC AAY52200;
XX
XX DT 14-MAR-2000 (first entry)
XX
XX DE Human la autoantigen peptide (LAP).
XX
XX KM La autoantigen; LAP, internal ribosome entry site; IRES; translation;
XX viral replication; RNA; antiviral agent; picornavirus; flavivirus;
XX coronavirus; hepatitis virus; rhinovirus; coxsackie virus;
XX parainfluenza virus; poliovirus; encephalomyocarditis virus;
XX foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
XX vesicular stomatitis virus.
XX
XX OS Homo sapiens.
XX
XX PN WO9961613-A2.
XX
XX PD 02-DEC-1999.
XX
XX PF 21-MAY-1999; 99WO-US011281.
XX
XX PR 22-MAY-1998; 98US-0086527P.
XX
XX PA (REGC) UNIV CALIFORNIA.
XX
XX PI Das S, Dasgupta A;
XX
XX DR WPI; 2000-062712/05.
XX
XX PT New yeast inhibitory peptide useful for inhibiting viral protein
XX translation and replication.
XX
XX PS Claim 5; Page 57; 81pp; English.
XX
XX This sequence is the la autoantigen binding domain (LAP). LAP is a
XX cellular protein which binds to an internal ribosome entry site (IRES).
XX The peptide is used to inhibit mRNA translation, viral protein
XX translation or viral replication. Viral mRNA translation is initiated at
XX the IRES and the LAP peptide prevents translation initiation factors from
XX binding at the site. The peptide can be used alone or in combination with
XX an inhibitor RNA (RNA see A4245200). The LAP peptide is useful as an
XX antiviral agent, which works through the inhibition of mRNA translation,
XX especially viral mRNA. Examples of viruses which can be inhibited are
XX picornavirus, flavivirus, coronavirus, hepatitis A B or C viruses,
XX rhinovirus, adenovirus, and parainfluenza virus, poliovirus, rhinovirus,

Query Match 100.0%; Score 84; DB 6; Length 16;
 Best Local Similarity 100.0%; Pred. No. 7.1e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OERAIRROVEYYFGDP 16
 |||||
 1 OERAIRROVEYYFGDP 16

RESULT 2

ABG72104
 ID ABG72104 standard; protein; 390 AA.

XX ABG72104;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 22740.

KW Drosophila; developmental biology; cell signalling; insecticide;
 pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PT 11-JUL-2000; 2000US-00614150.

PS (PEKS) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EM;

DR WPI; 2001-656860/75.

XX N-PSDB; ABL09419.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.

PS Disclosure; SEQ ID NO 22740; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABG57737-
 CC ABG72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 390 AA;

Query Match 92.9%; Score 78; DB 4; Length 390;
 Best Local Similarity 100.0%; Pred. No. 3.3e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OERAIRROVEYYFGD 15
 |||||
 50 OERAIRROVEYYFGD 64

RESULT 3

ABG72104
 ID ABG72104 standard; peptide; 18 AA.

AC ABG72104;

DT 28-JAN-2003 (first entry)

DE Viral replication inhibiting peptide, 761.

KW Viral replication inhibitor; IRES initiated translation; LAP;

KW internal ribosome entry site initiated translation; La antigen protein;
 KW viral life cycle; antiviral therapy; pharmaceutical; veterinary;

KW agricultural; horticultural; virucide.

OS Unidentified.

PN WO200283858-A2.

PD 24-OCT-2002.

PF 12-APR-2002; 2002WO-US011589.

PR 16-APR-2001; 2001US-00836073.

PA (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.

PI Dasgupta A, Das S, Baidya N;

DR WPI; 2003-058634/05.

PT New compound containing acidic and aromatic amino acids, useful as
 PT antiviral therapy in pharmaceutical, veterinary or
 PT agricultural/horticultural applications.

PS Claim 10; Page 16; 19pp; English.

CC The present invention relates to peptides and methods of inhibiting the
 CC replication of viruses that utilise internal ribosome entry site (IRES)
 CC initiated translation, and/or inhibiting viruses that utilise the La
 CC antigen protein (LAP) in any phase of their life cycle. The peptides of
 CC the invention compete with LAP and inhibit the utilisation of various
 CC biochemical and physiological functions of LAP required for a productive
 CC life cycle. The methods and compositions are useful as antiviral therapy
 CC in pharmaceutical, veterinary or agricultural/horticultural applications.
 CC ABG72101-ABG72119 represent peptides useful as antiviral agents

XX SQ Sequence 18 AA;

Query Match 73.8%; Score 62; DB 6; Length 18;
 Best Local Similarity 68.8%; Pred. No. 0.00063;

Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 OERAIRROVEYYFGDP 16
 |||||
 3 OERAIRROVEYYFGDP 18

RESULT 4

ABG72102
 ID ABG72102 standard; peptide; 18 AA.

XX ABG72102;

DT 28-JAN-2003 (first entry)

DE Viral replication inhibiting peptide, 702.

KW Viral replication inhibitor; IRES initiated translation; LAP;

KW internal ribosome entry site initiated translation; La antigen protein;
 KW viral life cycle; antiviral therapy; pharmaceutical; veterinary;

KW agricultural; horticultural; virucide.

OS Unidentified.

PN WO200283858-A2.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 16:58:20 ; Search time 50.0559 Seconds
(without alignments)
90.314 Million cell updates/sec

Title: US-09-836-073-19

Perfect score: 84
Sequence: 1 QERAIIRQVEYRQDF 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp19980s:*\n2: geneseqp19998s:*\n3: geneseqp2000s:*\n4: geneseqp2001s:*\n5: geneseqp2002s:*\n6: geneseqp2003as:*\n7: geneseqp2003bs:*\n8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	100.0	16	ABG72119	Abg72119 Viral rep
2	78	92.9	390	ABB65316	Abb65316 Drosophila
3	62	73.8	18	ABG72104	Abg72104 Viral rep
4	58	69.0	18	ABG72102	Abg72102 Viral rep
5	57	67.9	17	ABG72113	Abg72113 Viral rep
6	57	67.9	18	AAV52200	AAV52200 Human Ia
7	57	67.9	18	ABG72101	Abg72101 Viral rep
8	57	67.9	18	ABG72105	Abg72105 Viral rep
9	57	67.9	18	ABG72114	Abg72114 Viral rep
10	57	67.9	92	AAW03716	AAW03716 Human sec
11	57	67.9	408	AAW01351	AAW01351 Human aut
12	57	67.9	408	ABP65252	ABP65252 Hypoxia-r
13	57	67.9	408	ABP65391	ABP65391 Human pro
14	57	67.9	408	ADD46272	ADD46272 Human pro
15	57	67.9	408	ADP63955	ADP63955 Human pro
16	57	67.9	408	ADP62859	ADP62859 Human pro
17	57	67.9	415	ADP63989	ADP63989 Rat Prote
18	57	67.9	415	ADP63993	ADP63993 Rat Prote
19	57	67.9	439	ABG08417	Abg08417 Novel hum
20	57	67.9	460	AAH58987	AAH58987 Breast an
21	57	67.9	460	ABP41511	ABP41511 Human ova
22	54	64.3	18	ABG72115	Abg72115 Viral rep
23	54	64.3	21	AAH43394	AAH43394 La/SSB ep
24	53	63.1	411	ABP70526	ABP70526 Histone d
25	51	60.7	18	ABG72109	Abg72109 Viral rep

26	50	59.5	18	6	ABG72117	Abg72117 Viral rep
27	49	58.3	18	6	ABG72111	Abg72111 Viral rep
28	49	58.3	18	6	ABG72112	Abg72112 Viral rep
29	48	57.1	18	6	ABG72110	Abg72110 Viral rep
30	48	57.1	483	3	AAH29675	AAH29675 Arabidops
31	47	56.0	913	3	AAH47714	AAH47714 Arabidops
32	47	56.0	923	3	AAH47713	AAH47713 Arabidops
33	47	56.0	993	3	AAH47712	AAH47712 Arabidops
34	47	56.0	1379	4	ABP68940	ABP68940 Drosophila
35	46.5	55.4	19	6	ABG72116	Abg72116 Viral rep
36	45	53.6	18	6	ABG72103	Abg72103 Viral rep
37	44.5	53.0	18	6	ABG72118	Abg72118 Viral rep
38	44	52.4	315	7	ADP62048	ADP62048 Rat Prote
39	44	52.4	315	7	ADP62044	ADP62044 Rat Prote
40	44	52.4	315	7	ADP62044	ADP62044 Rat Prote
41	44	52.4	315	7	ADP62044	ADP62044 Rat Prote
42	43	51.2	466	6	ABP52760	ABP52760 Protein s
43	43	51.2	979	7	ADP64292	ADP64292 Human pro
44	42	50.0	18	6	ABG72107	Abg72107 Viral rep
45	42	50.0	275	6	ABP52759	ABP52759 Protein s

ALIGNMENTS

RESULT 1
ABG72119
ID ABG72119 standard; peptide: 16 AA.

XX ABG72119;

DT 28-JAN-2003 (first entry)

DE Viral replication inhibiting peptide, DROSOPHILA.

XX Viral replication inhibitor; IRES initiated translation; LAP; internal ribosome entry site initiated translation; La antigen protein; viral life cycle; antiviral therapy; pharmaceutical; veterinary;

KW agricultural; horticultural; virocid.

XX Drosophila sp.

OS WO200283858-A2.

XX 24-OCT-2002.

XX 12-APR-2002; 2002WO-US011589.

XX 16-APR-2001; 2001US-00836073.

XX (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.

XX Daegupye A, Das S, Baidya N;

XX WPI, 2003-058634/05.

XX New compound containing acidic and aromatic amino acids, useful as antiviral therapy in pharmaceutical, veterinary or agricultural/horticultural applications.

XX Claim 1, Page 15, 19pp; English.

XX The present invention relates to peptides and methods of inhibiting the replication of viruses that utilize internal ribosome entry site (IRES) initiated translation, and/or inhibiting viruses that utilize the La antigen protein (LAP) in any phase of their life cycle. The peptides of the invention compete with LAP and inhibit the utilization of various biochemical and physiological functions of LAP required for a productive life cycle. The methods and compositions are useful as antiviral therapy in pharmaceutical, veterinary or agricultural/horticultural applications.

XX ABG72101-ABG72119 represent peptides useful as antiviral agents

XX Sequence 16 AA:

XX

XX

XX

XX

This Page Blank (uspro)

Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 SKLEASTIROEYFGDA 18
|||||:|

Db 589 SKLEASTTKYADVFGNA 605

RESULT 15

US-10-651-183-20

; Sequence 20, Application US/10651183

; Publication No. US20040096929A1

; GENERAL INFORMATION:

; APPLICANT: KAWASAKI, GLENN

; APPLICANT: WEBB, HEATHER K.

; APPLICANT: OWENS, JEFFREY

; APPLICANT: LIEDTKE, RAYMOND

; APPLICANT: FOREST, DOREEN

; APPLICANT: LEGAZ, MARK

; APPLICANT: LAWSON, SOBOMABO

; TITLE OF INVENTION: ENZYMATIC CYCLING ASSAYS FOR HOMOCYSTEINE AND CYSTATIONINE

; FILE REFERENCE: 30865

; CURRENT APPLICATION NUMBER: US/10/651,183

; CURRENT FILING DATE: 2003-08-28

; PRIOR APPLICATION NUMBER: US/10/012,762

; PRIOR FILING DATE: 2002-09-26

; PRIOR APPLICATION NUMBER: 60/163,126

; PRIOR FILING DATE: 1999-11-02

; PRIOR APPLICATION NUMBER: 09/704,036

; PRIOR FILING DATE: 2000-11-01

; PRIOR APPLICATION NUMBER: 60/203,349

; PRIOR FILING DATE: 2000-05-10

; NUMBER OF SEQ ID NOS: 21

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 20

; LENGTH: 1252

; TYPE: PRT

; ORGANISM: Saccharomyces cerevisiae

; US-10-651-183-20

Query Match 47.8%; Score 43; DB 16; Length 1252;

Best Local Similarity 58.8%; Pred. No. 2.3e+02;

Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 SKLEASTIROEYFGDA 18
|||||:|

Db 589 SKLEASTTKYADVFGNA 605

Search completed: September 10, 2004, 18:11:57
Job time : 40.424 secs

Qy 1 VSKLEASTIRQEVYFG 16
Db 61 ISKGNLGRIRKEFYFG 76

RESULT 11

US-09-734-017A-32
; Sequence 32, Application US/09734017A
; Patent No. US2002014242A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Benz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reindl, Andreas
; APPLICANT: Clippus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Annette
; APPLICANT: Duwenig, Elke
; APPLICANT: Schmidt, Ralf-Michael
; TITLE OF INVENTION: Moss genes from *Physcomitrella patens* encoding proteins involved
; TITLE OF INVENTION: the
; TITLE OF INVENTION: synthesis of amino acids, vitamins, cofactors, nucleosides and
; FILE REFERENCE: BASF-NAE-1331-99-US
; CURRENT APPLICATION NUMBER: US/09/734,017A
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/171,100
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect
; SEQ ID NO 32
; LENGTH: 184
; TYPE: PRT
; ORGANISM: *Physcomitrella patens*
US-09-734-017A-32

Query Match 47.8%; Score 43; DB 9; Length 184;
Best Local Similarity 41.2%; Pred. No. 26;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 VSKLEASTIRQEVYFGD 17
Db 35 IHSAAQTMLEBYIGD 51

RESULT 12

US-10-464-811-20
; Sequence 20, Application US/10464811
; Publication No. US2004003319A1
; GENERAL INFORMATION:
; APPLICANT: Kraus, Jan
; APPLICANT: Oliveriusova, Jana
; TITLE OF INVENTION: Human Cytachionine B-Synthase Variants and Methods of Production
; FILE REFERENCE: 2848-49
; CURRENT APPLICATION NUMBER: US/10/464,811
; PRIOR FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: 60/389,541
; PRIOR FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 507
; TYPE: PRT
; ORGANISM: *Saccharomyces cerevisiae*
US-10-464-811-20

Query Match 47.8%; Score 43; DB 12; Length 507;
Best Local Similarity 58.8%; Pred. No. 82;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 SKLEASTIRQEVYFGDA 18

Db 351 SKLEASTIRQEVYFGDA 367

RESULT 13

US-10-451-467A-182
; Sequence 182, Application US/10451467A
; Publication No. US20040161840A1
; GENERAL INFORMATION:
; APPLICANT: CONTRERAS, ROLAND HENRI
; APPLICANT: EBERHARDT, INES
; APPLICANT: LUYTEN, WALTER HERMAN MARIA LOUIS
; APPLICANT: BEEKMAN, RIKA JOSEPHINA
; TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN
; TITLE OF INVENTION: YEAST AND FUNGI
; FILE REFERENCE: JAB-1667
; CURRENT APPLICATION NUMBER: US/10/451,467A
; PRIOR FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: EP 00870318.3
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: EP 01870002.1
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: EP 01870003.9
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 732
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 182
; LENGTH: 507
; TYPE: PRT
; ORGANISM: *Saccharomyces cerevisiae*
US-10-451-467A-182

Query Match 47.8%; Score 43; DB 16; Length 507;
Best Local Similarity 58.8%; Pred. No. 82;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 SKLEASTIRQEVYFGDA 18
Db 351 SKLEASTIRQEVYFGDA 367

RESULT 14

US-10-012-762-20
; Sequence 20, Application US/10012762
; Publication No. US20030138872A1
; GENERAL INFORMATION:
; APPLICANT: KAWASAKI, GLENN
; APPLICANT: WEBB, HEATHER K.
; APPLICANT: OWENS, JEFFREY
; APPLICANT: LIEDTKE, RAYMOND
; APPLICANT: FOREST, DOREEN
; APPLICANT: LEGAZ, MARK
; TITLE OF INVENTION: ENZYMATIC CYCLING ASSAYS FOR HOMOCYSTEINE AND CYSTATIONINE
; FILE REFERENCE: 30865
; CURRENT APPLICATION NUMBER: US/10/012,762
; PRIOR FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: 60/163,126
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: 09/704,036
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/203,349
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 1252
; TYPE: PRT
; ORGANISM: *Saccharomyces cerevisiae*
US-10-012-762-20

Query Match 47.8%; Score 43; DB 14; Length 1252;
Best Local Similarity 58.8%; Pred. No. 2.3e+02;

TYPE: PRT
ORGANISM: Homo sapiens
US-10-102-806-695

Query Match
Best Local Similarity 48.3%; Score 43.5; DB 14; Length 460;
Matches 10; Conservative 55.6%; Pred. No. 60; Mismatches 2; Indels 5; Gaps 1;

QY 1 VSKLEASTIRQ-EYRFGD 17
DB 62 MALEAKICHQIEYRFGD 79

RESULT 7
US-10-264-049-2643
Sequence 2643, Application US/10264049
Publication No. US20040005579A1
GENERAL INFORMATION:
APPLICANT: Birse et al.
FILE REFERENCE: Pal33pi
CURRENT APPLICATION NUMBER: US/10/264,049
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/18569
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: US 60/209,467
PRIOR FILING DATE: 2000-06-07
NUMBER OF SEQ ID NOS: 4360
SOFTWARE: Patentin Ver. 3.1
SEQ ID NO 2643
LENGTH: 460
TYPE: PRT
ORGANISM: Homo sapiens
US-10-264-049-2643

Query Match
Best Local Similarity 48.3%; Score 43.5; DB 15; Length 460;
Matches 10; Conservative 55.6%; Pred. No. 60; Mismatches 2; Indels 5; Gaps 1;

QY 1 VSKLEASTIRQ-EYRFGD 17
DB 62 MALEAKICHQIEYRFGD 79

RESULT 8
US-10-437-963-128214
Sequence 128214, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barakuk, Brad
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 128214
LENGTH: 169
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(169)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_3058C.1.pep

US-10-437-963-128214

Query Match
Best Local Similarity 47.8%; Score 43; DB 16; Length 169;
Matches 8; Conservative 30.8%; Pred. No. 24; Mismatches 7; Indels 1; Gaps 1;

QY 3 KLEASTIROE-----YRFGDA 18
DB 95 KLDITTVKQDELDKTLVHTYYGGA 120

RESULT 9
US-10-425-114-37684
Sequence 37684, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 37684
LENGTH: 172
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700337382_FLI.pep
US-10-425-114-37684

Query Match
Best Local Similarity 47.8%; Score 43; DB 12; Length 172;
Matches 8; Conservative 50.0%; Pred. No. 24; Mismatches 3; Indels 5; Gaps 0;

QY 1 VSKLEASTIROEYRFG 16
DB 61 ISKXNLGRIRKEFYFG 76

RESULT 10
US-10-425-114-71728
Sequence 71728, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 71728
LENGTH: 172
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: UC-ZMROMO17113D10_FLI.pep
US-10-425-114-71728

Query Match
Best Local Similarity 47.8%; Score 43; DB 12; Length 172;
Matches 8; Conservative 50.0%; Pred. No. 24; Mismatches 3; Indels 5; Gaps 0;

```

; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 19
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila
US-09-836-073-19

Query Match          49.4%; Score 44.5; DB 9; Length 16;
Best Local Similarity 71.4%; Pred. No. 0.91;
Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY      5 EASTIRQ-EYFQD 17
        | : ||| |||||
        2 ERAIRQVEYFQD 15

Db

RESULT 3
US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match          48.3%; Score 43.5; DB 9; Length 18;
Best Local Similarity 66.7%; Pred. No. 1.5;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY      4 LEASTIRQ-EYFQD 17
        ||| : |||||
        3 LEAQICQIEYFQD 17

Db

RESULT 4
US-10-170-385-477
; Sequence 477, Application US/10170385
; Publication No. US2003020372A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingsman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 532682000100
; CURRENT APPLICATION NUMBER: US/10/170,385
```

```

; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 477
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-170-385-477

Query Match          48.3%; Score 43.5; DB 12; Length 408;
Best Local Similarity 55.6%; Pred. No. 53;
Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY      1 VSKLEASTIRQ-EYFQD 17
        :: ||| |||||
        10 MALEAKICQIEYFQD 27

Db

RESULT 5
US-09-925-298-695
; Sequence 695, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 695
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-298-695

Query Match          48.3%; Score 43.5; DB 12; Length 460;
Best Local Similarity 55.6%; Pred. No. 60;
Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY      1 VSKLEASTIRQ-EYFQD 17
        :: ||| |||||
        62 MALEAKICQIEYFQD 79

Db

RESULT 6
US-10-102-806-695
; Sequence 695, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 695
; LENGTH: 460
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: September 10, 2004, 17:52:06 ; Search time 40.324 Seconds
(without alignments)
143.151 Million cell updates/sec

Title: US-09-836-073-18
Perfect score: 90
Sequence: 1 VSKLEASTIROEYFGDA 18

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1335176 seqs, 320689617 residues
Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppa/US06_NEW_PUB.pep:*
3: /cgn2_6/prodata/1/pubppa/US06_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*
7: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
12: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
13: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
16: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/prodata/1/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/prodata/1/pubppa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	90	100.0	18	9 US-09-836-073-18	Sequence 18, Appl
2	44.5	49.4	16	9 US-09-836-073-19	Sequence 19, Appl
3	43.5	48.3	18	9 US-09-836-073-2	Sequence 2, Appl
4	43.5	48.3	408	12 US-10-170-385-477	Sequence 477, App
5	43.5	48.3	460	12 US-09-925-298-695	Sequence 695, App
6	43.5	48.3	460	14 US-10-102-806-695	Sequence 695, App
7	43.5	48.3	169	15 US-10-264-049-2643	Sequence 2643, Ap
8	43	47.8	169	16 US-10-437-963-128214	Sequence 128214, A
9	43	47.8	172	12 US-10-425-114-37684	Sequence 37684, A
10	43	47.8	172	12 US-10-425-114-71728	Sequence 71728, A
11	43	47.8	184	9 US-09-734-017A-32	Sequence 32, Appl
12	43	47.8	507	12 US-10-464-811-20	Sequence 20, Appl
13	43	47.8	507	16 US-10-451-467A-182	Sequence 182, Appl
14	43	47.8	1252	14 US-10-012-762-20	Sequence 20, Appl
15	43	47.8	1252	16 US-10-651-183-20	Sequence 20, Appl

16	42.5	47.2	17	9 US-09-836-073-13	Sequence 13, Appl
17	42.5	47.2	18	9 US-09-836-073-1	Sequence 1, Appl
18	42.5	47.2	18	9 US-09-836-073-9	Sequence 9, Appl
19	42.5	47.2	18	9 US-09-836-073-14	Sequence 14, Appl
20	42.5	47.2	39	9 US-09-843-676-26	Sequence 26, Appl
21	42.5	47.2	39	9 US-09-766-253-26	Sequence 26, Appl
22	42.5	47.2	39	10 US-09-438-486-26	Sequence 26, Appl
23	42.5	47.2	39	12 US-10-325-810-216	Sequence 216, App
24	42.5	47.2	39	14 US-10-053-758-26	Sequence 26, Appl
25	42.5	47.2	39	14 US-10-054-295-26	Sequence 26, Appl
26	42.5	47.2	39	14 US-10-054-611-26	Sequence 26, Appl
27	42	46.7	19	9 US-09-836-073-16	Sequence 16, Appl
28	42	46.7	37	9 US-09-843-676-24	Sequence 24, Appl
29	42	46.7	37	9 US-09-766-253-24	Sequence 24, Appl
30	42	46.7	37	10 US-09-438-486-24	Sequence 24, Appl
31	42	46.7	37	14 US-10-053-758-24	Sequence 24, Appl
32	42	46.7	37	14 US-10-054-295-24	Sequence 24, Appl
33	42	46.7	37	14 US-10-054-611-24	Sequence 24, Appl
34	42	46.7	39	14 US-10-363-493-8277	Sequence 8277, Ap
35	42	46.7	895	16 US-10-437-963-147678	Sequence 147678, A
36	41	45.6	558	15 US-10-369-493-3398	Sequence 3398, Ap
37	41	45.6	1036	12 US-10-282-132A-47032	Sequence 47032, A
38	40.5	45.0	411	14 US-10-177-478-8	Sequence 8, Appl
39	40	44.4	546	15 US-10-369-493-1968	Sequence 1968, Ap
40	40	44.4	571	12 US-10-282-122A-44539	Sequence 44539, A
41	40	44.4	681	12 US-10-425-114-69655	Sequence 69655, A
42	40	44.4	768	12 US-10-425-114-68939	Sequence 68939, A
43	40	44.4	815	12 US-10-425-114-57809	Sequence 57809, A
44	40	44.4	876	9 US-09-815-242-5432	Sequence 5432, Ap
45	40	44.4	876	12 US-10-282-122A-43838	Sequence 43838, A

ALIGNMENTS

```
US-09-836-073-18
; Sequence 18, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836, 073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316, 630
; PRIORITY FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Mosquito
US-09-836-073-18

Query Match      100.0%; Score 90; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSKLEASTIROEYFGDA 18
Db 1 VSKLEASTIROEYFGDA 18

RESULT 2
US-09-836-073-19
; Sequence 19, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
```

This Page Blank (uspto)

CORRESPONDENCE ADDRESS:

ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER: US/09/430,323

FILING DATE: 29-Oct-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US 08/724,643

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002930US

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 37 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 24:

US-09-430-323-24

Query Match 46.7%; Score 42; DB 4; Length 37;

Best Local Similarity 100.0%; Pred. No. 1;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 QERYFGD 17

Db 4 QERYFGD 10

RESULT 15

US-09-328-352-8030

Sequence 8030, Application US/09328352

Patent No. 6562958

GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

FILE REFERENCE: GTC99-03PA

CURRENT APPLICATION NUMBER: US/09/328,352

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 8252

SEQ ID NO 8030

LENGTH: 287

TYPE: PRT

ORGANISM: Acinetobacter baumannii

US-09-328-352-8030

Query Match 44.4%; Score 40; DB 4; Length 287;

Best Local Similarity 57.1%; Pred. No. 26;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 SKLEASTIRQERYF 15

Db 170 SKQERITTKDRYF 183

Search completed: September 10, 2004, 18:05:12
Job time: 16.8883 secs

APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin H.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809e1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-24

Query Match 46.7%; Score 42; DB 3; Length 37;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 OEYFGD 17
| | | | |
4 OEYFGD 10

Db 4 OEYFGD 10

RESULT 13
US-08-854-050-24
Sequence 24, Application US/08854050
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin H.
APPLICANT: Andrews, William H.

TITLE OF INVENTION: No. 6261836e1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-24

Query Match 46.7%; Score 42; DB 3; Length 37;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 OEYFGD 17
| | | | |
4 OEYFGD 10

Db 4 OEYFGD 10

RESULT 14
US-09-430-323-24
Sequence 24, Application US/09430323
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin H.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6309867e1 Telomerase
NUMBER OF SEQUENCES: 225

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,181B
FILING DATE: 29-Sep-1997
CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Aussenhus, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 216:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 216:
US-09-402-181B-216

Query Match 47.2%; Score 42.5; DB 4; Length 39;
Best Local Similarity 81.8%; Pred. No. 0.91;
Matches 9; Conservative 1; Mismatches 0; Indels 1; Gaps 1

QY 9 IRQ-EYRFGDA 18
: || || || || ||
Db 2 LROVEYRFGDA 12

RESULT 11
US-09-721-456-216
; Sequence 216, Application US/09721456
; Patent No. 6617110
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H

```

```

TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/721,456
FILING DATE: 22-NO. 6617110-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997

ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 216:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 216:
US-09-721-456-216

Query Match 47.2% Score 42.5; DB 4, Length 39,
Best Local Similarity 81.8%; Pred.No. 0.91;
Matches 9; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

CY 9 IRQ-EYTFGDA 18
:|||||
Db 2 LROVEYTFGDA 12

RESULT 12
US-08-851-843A-24
; Sequence 24, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.

```

```

1      Chapman, Karen B.
2      Morin, Gregg B.
3      Harley, Calvin
4      Andrews, William H.
5      TITLE OF INVENTION: No. 6309867e1 Telomerase
6      NUMBER OF SEQUENCES: 225
7      CORRESPONDENCE ADDRESS:
8      ADDRESSEE: Townsend and Townsend and Crew LLP
9      STREET: Two Embarcadero Center, 8th Floor
10     CITY: San Francisco
11     STATE: California
12     COUNTRY: United States of America
13     ZIP: 94111
14
15     COMPUTER READABLE FORM:
16     MEDIUM TYPE: Floppy disk
17     COMPUTER: IBM PC compatible
18     OPERATING SYSTEM: PC-DOS/MS-DOS
19     SOFTWARE: Patent Release #1.0, Version #1.30
20     CURRENT APPLICATION DATA:
21     APPLICATION NUMBER: US/09/430,323
22     FILING DATE: 29-Oct-1999
23     CLASSIFICATION: <Unknown>
24     PRIOR APPLICATION DATA:
25     APPLICATION NUMBER: US 08/854,050
26     FILING DATE: 09-MAY-1997
27     APPLICATION NUMBER: US 08/851,843
28     FILING DATE: 06-MAY-1997
29     APPLICATION NUMBER: US 08/846,017
30     FILING DATE: 25-APR-1997
31     APPLICATION NUMBER: US 08/844,419
32     FILING DATE: 18-APR-1997
33     APPLICATION NUMBER: US 08/724,643
34     FILING DATE: 01-OCT-1996
35     ATTORNEY/AGENT INFORMATION:
36     NAME: Apple, Randolph T.
37     REGISTRATION NUMBER: 36,429
38     REFERENCE/DOCKET NUMBER: 015389-002930US
39     TELECOMMUNICATION INFORMATION:
40     TELEPHONE: (415) 576-0200
41     TELEFAX: (415) 576-0300
42     INFORMATION FOR SEQ ID NO: 26:
43     SEQUENCE CHARACTERISTICS:
44     LENGTH: 39 amino acids
45     TYPE: amino acid
46     STRANDEDNESS: not relevant
47     TOPOLOGY: not relevant
48     MOLECULE TYPE: peptide
49     SEQUENCE DESCRIPTION: SEQ ID NO: 26:
50
51     US-09-430-323-26
52
53     Query Match          47.2%; Score 42.5; DB 4; Length 39;
54     Best Local Similarity 81.8%; Pred.No.0.91;
55     Matches          9; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
56
57     QY      9 IRQ-EYTFGDA 18
58           :||| |||||
59           2 LRQVEYTFGDA 12
60
61     RESULT 10
62     US-09-402-181B-216
63     ; Sequence 216, Application US/09402181B
64     ; Patent No. 6610839
65     ; GENERAL INFORMATION:
66     ; APPLICANT: Cecb, Thomas R.
67     ;               Lingner, Joachim
68     ;               Nakamura, Toru
69     ;               Chapman, Karen B.
70     ;               Morin, Gregg B.
71     ;               Harley, Calvin B.
72     ;               Andrews, William H.
73     ; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
74     ; NUMBER OF SEQUENCES: 633

```

TITLE OF INVENTION: No. 6093809e1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ. ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-26
Query Match 47.2%; Score 42.5; DB 3; Length 39;
Best Local Similarity 81.8%; Pred. No. 0.91;
Matches 9; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 9 IRO-EYFFGDA 18
:|||||
Db 2 IROVEYFFGDA 12

RESULT 7
US-08-974-549A-216
Sequence 216, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Hatley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco

STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ. ID NO: 216:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-216
Query Match 47.2%; Score 42.5; DB 3; Length 39;
Best Local Similarity 81.8%; Pred. No. 0.91;
Matches 9; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 9 IRO-EYFFGDA 18
:|||||
Db 2 IROVEYFFGDA 12

RESULT 8
US-08-854-050-26
Sequence 26, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.

```
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/203,349
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 1252
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-012-762-20

Query Match
Best Local Similarity 47.8%; Score 43; DB 4; Length 1252;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 SKLEASTIRQRYFGDA 18
Db 589 SKLEASTIRKYADVFGNA 605

RESULT 3
US-09-704-036B-20
; Sequence 20, Application US/09704036B
; Patent No. 6664073
; GENERAL INFORMATION:
; APPLICANT: KAMASAKI, GLENN
; APPLICANT: WEBB, HEATHER
; APPLICANT: OWENS, JEFFREY
; APPLICANT: LIEDTKE, RAYMOND
; APPLICANT: FOREST, DOREEN
; APPLICANT: LEGAZ, MARK
; APPLICANT: LAMSON, SOBOAMBO
; TITLE OF INVENTION: ENZYMATIC CYCLING ASSAYS FOR HOMOCYSTEINE AND CYSTATINONINE
; FILE REFERENCE: 30865
; CURRENT APPLICATION NUMBER: US/09/704,036B
; CURRENT FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/163,126
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: 09/704,036
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/203,349
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 1252
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-704-036B-20

Query Match
Best Local Similarity 47.8%; Score 43; DB 4; Length 1252;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 SKLEASTIRQRYFGDA 18
Db 589 SKLEASTIRKYADVFGNA 605

RESULT 4
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRIS-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
```

```
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3

Query Match
Best Local Similarity 47.2%; Score 42.5; DB 3; Length 18;
Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 4 LEASTIRQ-XYFGD 17
Db 3 LEAKICHOIEYFGD 17

RESULT 5
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRIS-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: B-LAP
US-09-316-630-4

Query Match
Best Local Similarity 47.2%; Score 42.5; DB 3; Length 18;
Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 4 LEASTIRQ-XYFGD 17
Db 3 LEAKICHOIEYFGD 17

RESULT 6
US-08-851-843A-26
; Sequence 26, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Langner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:43:01; Search time 15.8883 Seconds
(without alignments)
58.488 Million cell updates/sec

Title: US-09-836-073-18

Sequence: 1 VSKLEASTIRQRYFGDA 18

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B COMB.pep:*
3: /cgn2_6/prodata/2/iaa/6A COMB.pep:*
4: /cgn2_6/prodata/2/iaa/6B COMB.pep:*
5: /cgn2_6/prodata/2/iaa/PCTUS COMB.pep:*
6: /cgn2_6/prodata/2/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43.5	48.3	136	4 US-09-134-001C-4980	Sequence 4980, App
2	43	47.8	1252	4 US-10-012-762-20	Sequence 20, App1
3	43	47.8	1252	4 US-09-704-0368-20	Sequence 20, App1
4	42.5	47.2	18	3 US-09-316-630-3	Sequence 3, App1
5	42.5	47.2	18	3 US-09-316-630-4	Sequence 4, App1
6	42.5	47.2	39	3 US-08-851-843A-26	Sequence 26, App1
7	42.5	47.2	39	3 US-08-851-843A-26	Sequence 26, App1
8	42.5	47.2	39	3 US-08-854-050-26	Sequence 26, App1
9	42.5	47.2	39	4 US-09-430-323-26	Sequence 26, App1
10	42.5	47.2	39	4 US-09-402-181B-216	Sequence 216, App
11	42.5	47.2	39	4 US-09-721-456-216	Sequence 216, App
12	42	46.7	37	3 US-08-851-843A-24	Sequence 24, App1
13	42	46.7	37	3 US-08-854-050-24	Sequence 24, App1
14	42	46.7	37	4 US-09-430-323-24	Sequence 24, App1
15	40	44.4	287	4 US-09-328-352-8030	Sequence 8030, App
16	40	44.4	546	4 US-09-457-040B-24	Sequence 24, App1
17	40	44.4	876	1 US-08-785-071A-2	Sequence 2, App1
18	40	44.4	876	1 US-09-012-872-2	Sequence 2, App1
19	39.5	43.9	506	2 US-08-849-480A-5	Sequence 5, App1
20	39	43.3	418	4 US-09-328-352-5700	Sequence 5700, App
21	39	43.3	550	1 US-08-674-168-29	Sequence 29, App1
22	39	43.3	550	3 US-08-985-908-19	Sequence 19, App1
23	39	43.3	550	3 US-08-852-730-4	Sequence 4, App1
24	39	43.3	550	3 US-08-985-916-11	Sequence 11, App1
25	39	43.3	872	4 US-09-540-236-2346	Sequence 2346, App
26	39	43.3	872	4 US-09-198-452A-163	Sequence 163, App
27	38.5	42.8	-128	6 5514582-33	Patent No. 5514582

28	38	42.2	147	4 US-09-621-976-4310	Sequence 4310, App
29	38	42.2	226	4 US-09-489-039A-7662	Sequence 7662, App
30	38	42.2	423	1 US-08-844-064-7	Sequence 7, App1
31	38	42.2	423	3 US-09-009-433-7	Sequence 7, App1
32	38	42.2	477	1 US-08-136-922-2	Sequence 2, App1
33	38	42.2	655	4 US-08-556-422A-3	Sequence 3, App1
34	38	42.2	771	1 US-08-121-713D-54	Sequence 54, App1
35	38	42.2	771	1 US-08-835-268-54	Sequence 54, App1
36	38	42.2	771	2 US-09-060-692-54	Sequence 54, App1
37	38	42.2	771	3 US-08-833-391-54	Sequence 54, App1
38	38	42.2	771	4 US-09-060-610-54	Sequence 54, App1
39	38	42.2	771	5 PCT-US84-10151A-54	Sequence 54, App1
40	38	42.2	883	2 US-08-953-492-2	Sequence 2, App1
41	38	42.2	897	4 US-09-134-001C-3600	Sequence 3600, App
42	37.5	41.7	249	4 US-09-134-001C-3910	Sequence 3910, App
43	37	41.1	21	4 US-08-475-955-20	Sequence 20, App1
44	37	41.1	38	3 US-08-851-843A-25	Sequence 25, App1
45	37	41.1	38	3 US-08-974-549A-214	Sequence 214, App

ALIGNMENTS

```

RESULT 1
US-09-134-001C-4980
; Sequence 4980, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4980
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4980

Query Match      48.3%; Score 43.5; DB 4; Length 136;
Best Local Similarity 57.9%; Pred. No. 2.6;
Matches 11; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

Cy      1 VSKLEAST-IRQRYFGDA 18
Db      98 ISKLESTDERQRYFFDA 116

RESULT 2
US-10-012-762-20
; Sequence 20, Application US/10012762
; Patent No. 6635438
; GENERAL INFORMATION:
; APPLICANT: KAWASAKI, GLENN
; APPLICANT: WEBB, HEATHER K.
; APPLICANT: OWENS, JEFFREY
; APPLICANT: LIEBTE, RAYMOND
; APPLICANT: FOREST, DOREEN
; APPLICANT: LEGAZ, MARK
; APPLICANT: LAWSON, SOBOMABO
; TITLE OF INVENTION: ENZYMAIC CYCLING ASSAYS FOR HOMOCYSTEINE AND CYSTATHIONINE
; FILE REFERENCE: 30865
; CURRENT APPLICATION NUMBER: US/10/012,762
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: 60/163,126
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: 09/704,036

```

This Page Blank (uspto)

DE Rat Protein P38656, SEQ ID NO 9935.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 XX
 OS Rattus norvegicus.

XX WO2003016475-A2.

XX 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-033347P.

XX (GEHO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

DR MPI; 2003-268312/26.

XX GENBANK: P38656.

PT New composition comprising two or more isolated polypeptides, useful for
 XX preparing a medicament for treating pain in an animal.

PS Claim 1, Page; 1017P; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 415 AA:

Query Match 48.3%; Score 43.5; DB 7; Length 415;
 Best Local Similarity 55.6%; Pred. No. 50;
 Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 VSKLEASTIRO-EYFQD 17
 :: ||| |||||
 DB 10 MAALBAKICHQIYFQD 27

Search completed: September 10, 2004, 17:51:55
 Job time : 58.4128 secs

PD 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
PI MPI; 2003-268312/26.
DR GENBANK; P05455.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page: 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 408 AA;
SQ
Query Match 48.3%; Score 43.5; DB 7; Length 408;
Best Local Similarity 55.6%; Pred. No. 49;
Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
QY 1 VSKLEASTIRQ-EYFPGD 17
Db 10 MAALFAKICHOIEYFPGD 27
RESULT 14
ADE62859
ID ADE62859 standard; protein; 408 AA.
XX
XX ADE62859;
AC
XX
XX 29-JAN-2004 (first entry)
DT
XX
XX Human Protein P05455, SEQ ID NO 8793.
DE
XX
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.

XX
XX Homo sapiens.
OS
XX WO2003016475-A2.
EN
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
PI MPI; 2003-268312/26.
DR GENBANK; P05455.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page: 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 408 AA;
SQ
Query Match 48.3%; Score 43.5; DB 7; Length 408;
Best Local Similarity 55.6%; Pred. No. 49;
Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
QY 1 VSKLEASTIRQ-EYFPGD 17
Db 10 MAALFAKICHOIEYFPGD 27
RESULT 15
ADE63989
ID ADE63989 standard; protein; 415 AA.
XX
XX ADE63989;
AC
XX
XX 29-JAN-2004 (first entry)
DT
XX

XX Woolf C, D'urso D, Befort K, Costigan M;
 PI WPI; 2003-268312/26.
 DR GENBANK; P05455.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX
 PS Claim 1, Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 408 AA;

XX Query Match 48.3%; Score 43.5; DB 7; Length 408;
 XX Best Local Similarity 55.6%; Pred. No. 49;
 XX Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 VKLEASTIRQ-EYFGD 17
 ::||| |
 Db 10 MALEAKICHQIEYFGD 27

XX RESULT 12
 XX ADD46272
 XX ID ADD46272 standard; protein; 408 AA.

XX ADD46272;

XX 29-JAN-2004 (first entry)

XX Human Protein P05455, SEQ ID NO 11947.

XX Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PA (GENO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;
 PI WPI; 2003-268312/26.
 DR GENBANK; P05455.

XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.

XX Claim 1, Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 408 AA;

XX Query Match 48.3%; Score 43.5; DB 7; Length 408;
 XX Best Local Similarity 55.6%; Pred. No. 49;
 XX Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 VKLEASTIRQ-EYFGD 17
 ::||| |
 Db 10 MALEAKICHQIEYFGD 27

XX RESULT 13
 XX ADE63995
 XX ID ADE63995 standard; protein; 408 AA.

XX ADE63995;

XX 29-JAN-2004 (first entry)

XX Human Protein P05455, SEQ ID NO 9941.

XX Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

DE Human autoantigen Ia(SS-B).
XX
XX Autoimmune disease; La autoantigen; Sjogren's syndrome;
KW systemic lupus erythematosus; diagnosis.
XX
XX Homo sapiens.
XX
XX USS541291-A.
XX
XX 30-JUL-1996.
XX
XX 27-MAY-1987; 87US-00054871.
XX
XX 31-DEC-1984; 84US-00687908.
XX
XX (UYDU-) UNIV DUKE.
XX
XX Keene JD;
XX
XX WPI; 1996-362015/36.
XX
XX Auto-antigen U2-RNP, associated with severe polymyositis-scleroderma
PT overlap syndrome - useful for diagnosis and treatment of autoimmune
PT diseases.
XX
XX PS Disclosure; Col 15-16; 21pp; English.
XX
XX The human lupus antigen (La) is diagnostic for Sjogren's syndrome, as
CC well as occurring in systemic lupus erythematosus patients. The La protein
CC is clinically related to the Ro protein that is highly common among
CC autoimmune patients. La and Ro antigens sometimes reside on the same
CC cellular ribonucleoprotein particle, most La patients contain some Ro
CC antibodies and vice versa. La cDNA has been isolated from a human liver
CC library. (Updated on 25-MAR-2003 to correct PF field.)
XX
XX SQ Sequence 408 AA;

Query Match 48.3%; Score 43.5; DB 2; Length 408;
Best Local Similarity 55.6%; Pred. No. 49;
Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 VSKLEASTIRQ-EYFGD 17
::||| |
DB 10 MALEAKICHQIEYFGD 27

RESULT 10
ID ABP65252
AC ABP65252; standard; protein; 408 AA.
XX
XX 12-NOV-2002 (first entry)
XX
XX DE Hypoxia-regulated protein #126.
XX
XX Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;
KW antiinflammatory; vulnery; gynecological; ophthalmological; vaccine;
KW hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;
KW ischaemic condition; reperfusion injury; retinopathy; neonatal stress;
KW preclapmsia; atherosclerosis; inflammatory condition; wound healing;
KW inflammation; erythropoiesis; hair loss; human.
XX
XX OS Homo sapiens.
XX
XX WO200246465-A2.
XX
XX 13-JUN-2002.
XX
XX 10-DEC-2001; 2001WO-GB005458.
XX
XX 08-DEC-2000; 2000GB-00030076.
XX
XX 08-FEB-2001; 2001GB-00003156.
XX
XX PR

PR 25-OCT-2001; 2001GB-00025666.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX White J, Mundy CR, Ward NR, Krige D, Kingman SM, Harris RA;
PI Rayner WN;
XX
XX WPI; 2002-627238/67.
XX
XX Identifying a gene involved in disease for treating hypoxia-regulated
PT conditions, comprises comparing the transcriptome/proteome of two cell
PT types under different conditions and identifying a differentially
PT regulated gene.
XX
XX Claim 35; Page 425; 538pp; English.
XX
XX The present invention relates to methods for identifying genes and
CC proteins that are implicated in a specific disease or physiological
CC condition. The method comprises comparing the transcriptome/proteome of a
CC specialised cell type implicated in a disease or condition with that of a
CC second specialised cell type, under two experimental conditions, and
CC identifying a gene that is differentially regulated in the two
CC specialised cell types under experimental conditions. ABV77873-ABV78116
CC and ABP65061-ABP65257 were identified using the methods of the invention.
CC The coding sequences and proteins are useful for treating a disease in a
CC patient, for manufacture of a medicament for treating hypoxia-regulated
CC conditions, and for regulating tumorigenesis, angiogenesis, apoptosis,
CC biological response to hypoxia conditions, or hypoxic-associated
CC pathology in a patient. The coding sequences and proteins are also useful
CC for monitoring the therapeutic treatment of a disease or physiological
CC condition, such as cancer, ischaemic conditions, reperfusion injury,
CC retinopathy, neonatal stress, preclapmsia, atherosclerosis, inflammatory
CC conditions, wound healing, inflammation, erythropoiesis or hair loss
XX
XX SQ Sequence 408 AA;

Query Match 48.3%; Score 43.5; DB 5; Length 408;
Best Local Similarity 55.6%; Pred. No. 49;
Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 VSKLEASTIRQ-EYFGD 17
::||| |
DB 10 MALEAKICHQIEYFGD 27

RESULT 11
ID ADE63991
AC ADE63991; standard; protein; 408 AA.
XX
XX 29-JAN-2004 (first entry)
XX
XX DE Human Protein P05455, SEQ ID NO 9937.
XX
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
XX OS Homo sapiens.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
XX
XX 01-NOV-2001; 2001US-0346382P.
XX
XX 26-NOV-2001; 2001US-0333347P.
XX
XX (GENO) GEN HOSPITAL CORP.
XX
XX (FARB) BAYER AG.

CC antigen protein (IAP) in any phase of their life cycle. The peptides of
 CC the invention compete with IAP and inhibit the utilisation of various
 CC biochemical and physiological functions of IAP required for a productive
 CC life cycle. The methods and compositions are useful as antiviral therapy
 CC in pharmaceutical, veterinary or agricultural/horticultural applications.
 CC ABG72101-ABG72119 represent peptides useful as antiviral agents

CC
 XX
 SQ Sequence 18 AA;

Query Match 48.3%; Score 43.5; DB 6; Length 18;
 Best Local Similarity 66.7%; Pred. No. 1.1;
 Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 4 LEASTIRQ-EYFEGD 17
 |||:|||||
 Db 3 LEAQCQQLRYFEGD 17

RESULT 7
 AAG01351

ID AAG01351 standard; protein; 92 AA.

AC AAG01351;

DT 06-OCT-2000 (first entry)

DE Human secreted protein, SEQ ID NO: 5432.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KW gene therapy; chromosome mapping.

XX Homo sapiens.

XX EPI033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-00200610.

XX 26-FEB-1999; 99US-0122487P.

PA (GSEST) GENSET.

DR Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

DR N-PSDB; AAC01357.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

PS Claim 13; SEQ ID NO 5432; 71pp + Sequence Listing; English.

XX
 CC The present sequence is a polypeptide encoded by one of a large number of
 CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
 CC prepared from total human RNAs or polyA+ RNAs derived from 30 different
 CC tissues. EST sequences usually correspond mainly to the 3' untranslated
 CC region (UTR) of the mRNA because they are often obtained from oligo-dT
 CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
 CC sequences derived from the 5' ends of mRNAs and even in those cases where
 CC longer cDNA sequences have been obtained, the full 5' UTR is rarely
 CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can
 CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
 CC are also used in diagnostic, forensic, gene therapy and chromosome
 CC mapping procedures. They are used to obtain upstream regulatory sequences
 CC and to design expression and secretion vectors

XX
 SQ Sequence 92 AA;

Query Match 48.3%; Score 43.5; DB 3; Length 92;
 Best Local Similarity 55.6%; Pred. No. 8;
 Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 VSKLEASTIRQ-EYFEGD 17
 ::|||:|||||
 Db 10 MALEAKICHQLEYFEGD 27

RESULT 8

ID ABP40135 standard; protein; 136 AA.

AC ABP40135;

DT 24-JUL-2002 (first entry)

DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:4980.

XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;

KW antibacterial; gene therapy.

XX Staphylococcus epidermidis.

XX US6380370-B1.

PD 30-APR-2002.

PF 13-AUG-1998; 98US-00134001.

XX 14-AUG-1997; 97US-0055779P.

PR 08-NOV-1997; 97US-0064964P.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Doucette-Stamm LA, Bush D;

XX WPI; 2002-381255/41.

DR N-PSDB; ABN92680.

PT Novel isolated nucleic acid encoding a Staphylococcus epidermis
 PT polypeptide, useful for diagnosing and treating bacterial infections.

XX Disclosure; SEQ ID NO 4980; 267pp; English.

XX
 CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
 CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
 CC antibacterial activity and can be used in gene therapy. The sequences can
 CC also be used in the diagnosis and treatment of bacterial infections,
 CC particularly S. epidermidis infections. The sequences can be used to
 CC screen for compounds able to interfere with the S. epidermidis life cycle
 CC or inhibit S. epidermidis infection. N.B. The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the USPTO web site

XX
 SQ Sequence 136 AA;

Query Match 48.3%; Score 43.5; DB 5; Length 136;
 Best Local Similarity 57.9%; Pred. No. 13;
 Matches 11; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 VSKLEAST-IRQLEYFEGD 18
 :|||:|||||
 Db 98 ISKLRSETDERQVYFEDA 116

RESULT 9

ID AAM03716 standard; protein; 408 AA.

AC AAM03716;

DT 25-MAR-2003 (revised)

XX 12-MAR-1997 (first entry)

```

PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

```

Query Match 48.9%; Score 44; DB 3; Length 413;
 Best Local Similarity 38.9%; Pred. No. 41;
 Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```

QY 1 VSKLEASTIROEYRGDA 18
    :|:|:|:|:|:|
DB 8 ILRCSSCTMRREYFPHDA 25

RESULT 5
AB57841
ID AB57841 standard; protein; 512 AA.
XX
AC AB57841;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 315.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
  pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE ) PE CORP NY.
XX

```

```

PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-6556860/75.
DR N-PSDB; ABL01944.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
  genes from Drosophila and for elucidating cell signalling and cell-cell
  interactions.
XX
PS Disclosure; SEQ ID NO 315; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
  capable of detecting 1000 or more genes from Drosophila. The invention is
  useful in developmental biology and in elucidating cell signalling and
  cell-cell interactions in higher eukaryotes for the development of
  insecticides, therapeutics and pharmaceutical drugs. The invention
  discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
  sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-
  ABL82072). The sequence data for this patent did not form part of the
  printed specification, but was obtained in electronic format directly
  from WIPO at ftp.wipo.int/pub/published_pct_sequences
CC
XX
SQ Sequence 512 AA;

```

Query Match 48.9%; Score 44; DB 4; Length 512;
 Best Local Similarity 57.1%; Pred. No. 53;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 VSKLEASTIROEY 14
    :|:|:|:|:|
DB 278 ISPLQNAQIROSY 291

```

```

RESULT 6
ABG72102
ID ABG72102 standard; peptide; 18 AA.
XX
AC ABG72102;
XX
DT 28-JAN-2003 (first entry)
XX
DE Viral replication inhibiting peptide, 702.
XX

```

```

KW Viral replication inhibitor; IRES initiated translation; LAMP;
  internal ribosome entry site initiated translation; Ia antigen protein;
  viral life cycle; antiviral therapy; pharmaceutical; veterinary;
  agricultural; horticultural; virucide.
XX
OS Unidentified.
XX
PN WO200283858-A2.
XX
PD 24-OCT-2002.
XX
PF 12-APR-2002; 2002WO-US011589.
XX
PR 16-APR-2001; 2001US-00836073.
XX
PA (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.
XX
PI Dasgupta A, Das S, Balda N;
XX
DR WPI; 2003-058634/05.
XX

```

New compound containing acidic and aromatic amino acids, useful as
 antiviral therapy in pharmaceutical, veterinary or
 agricultural/horticultural applications.

Claim 10; Page 16; 19pp; English.

The present invention relates to peptides and methods of inhibiting the
 replication of viruses that utilise internal ribosome entry site (IRES)
 initiated translation, and/or inhibiting viruses that utilise the Ia

```

XX 25-FEB-2000; 2000EP-00301439.
PR 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 04-MAY-1999; 99US-0132048P.
PR 06-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 07-MAY-1999; 99US-0132486P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134220P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134841P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140654P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-014091P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142380P.

PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-014576P.
PR 27-JUL-1999; 99US-0145813P.
PR 27-JUL-1999; 99US-0145818P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147933P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0150566P.
PR 25-AUG-1999; 99US-0150884P.
PR 26-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-015130P.
PR 01-SEP-1999; 99US-0152363P.
PR 07-SEP-1999; 99US-0153070P.
PR 10-SEP-1999; 99US-0153758P.
PR 13-SEP-1999; 99US-0154018P.
PR 15-SEP-1999; 99US-0154039P.
PR 16-SEP-1999; 99US-0154779P.
PR 20-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.

```

Query Match 100.0%; Score 90; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 4.6e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSKLEASTIRQVEYFGDA 18
 |||||
 DB 1 VSKLEASTIRQVEYFGDA 18

RESULT 2

ABB65316
 ID ABB65316 standard; protein; 390 AA.

XX ABB65316;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 22740.

XX Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656660/75.

XX N-PSDB; ABL09419.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signaling and cell-cell

XX interactions.

XX Dislosure; SEQ ID NO 22740; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signaling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (AB116176-AB130511), expressed DNA

XX sequences (AB101840-AB146175) and the encoded proteins (ABB57737-

XX ABB72072). The sequence data for this patent did not form part of the

XX printed specification, but was obtained in electronic format directly

XX from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 390 AA;

Query Match 58.3%; Score 52.5; DB 4; Length 390;

Best Local Similarity 66.7%; Pred. No. 1.1;

Matches 12; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

OY 2 SKLEASTIRQ-EYFGDA 18
 :|||:|||||
 DB 48 TKQERAILIRQVEYFGDA 65

RESULT 3

ABG72119
 ID ABG72119 standard; peptide; 16 AA.

AC ABG72119;

XX 28-JAN-2003 (first entry)

XX Viral replication inhibiting peptide, DROSOPHILA.

XX Viral replication inhibitor; IRES initiated translation; LAP;

XX internal ribosome entry site initiated translation; La antigen protein;

XX viral life cycle; antiviral therapy; pharmaceutical; veterinary;

XX agricultural; horticultural; virucide.

XX Drosophila sp.

XX WO200283858-A2.

XX 24-OCT-2002.

XX 12-APR-2002; 2002WO-US011589.

XX 16-APR-2001; 2001US-00836073.

XX (UTCA-) UNIV CALIFORNIA OFFICE PRESIDENT.

XX Dasgupta A, Das S, Baidya N;

XX WPI; 2003-058634/05.

XX New compound containing acidic and aromatic amino acids, useful as

XX antiviral therapy in pharmaceutical, veterinary or

XX agricultural/horticultural applications.

XX Claim 1; Page 15; 19pp; English.

XX The present invention relates to peptides and methods of inhibiting the

XX replication of viruses that utilise internal ribosome entry site (IRES)

XX initiated translation, and/or inhibiting viruses that utilise the La

XX antigen protein (LAP) in any phase of their life cycle. The peptides of

XX the invention compete with LAP and inhibit the utilisation of various

XX biochemical and physiological functions of LAP required for a productive

XX life cycle. The methods and compositions are useful as antiviral therapy

XX in pharmaceutical, veterinary or agricultural/horticultural applications.

XX ABG72101-ABG72119 represent peptide useful as antiviral agents

XX Sequence 16 AA;

XX Query Match 49.4%; Score 44.5; DB 6; Length 16;

XX Best Local Similarity 71.4%; Pred. No. 0.62;

XX Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

OY 5 EASTIRQ-EYFGD 17
 :|||:|||||
 DB 2 ERATIRQVEYFGD 15

RESULT 4

AAAG38952
 ID AAAG38952 standard; protein; 413 AA.

XX AAAG38952;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 48129.

XX Protein identification; signal transduction pathway; metabolic pathway;

XX hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 16:58:20 ; Search time 56.3129 Seconds
(without alignments)
90.314 Million cell updates/sec

Title: US-09-836-073-18

Perfect score: 90

Sequence: 1 VSKLEASTIRQRYFGDA 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	18	6	ABG72118 Viral rep
2	52.5	58.3	390	4	ABB65316 Drosophila
3	44.5	49.4	16	6	ABG72119 Viral rep
4	44	48.9	413	3	AAG38952 Arabidops
5	44	48.9	512	6	ABBS7841 Drosophila
6	43.5	48.3	18	6	ABG72102 Viral rep
7	43.5	48.3	92	3	AAG01351 Human sec
8	43.5	48.3	136	5	ABP40135 Staphyloc
9	43.5	48.3	408	5	AAW03716 Human aut
10	43.5	48.3	408	5	ABP65252 Hypoxia-r
11	43.5	48.3	408	7	ADBE63991 Human pro
12	43.5	48.3	408	7	ADDA6272 Human pro
13	43.5	48.3	408	7	ADBE63995 Human pro
14	43.5	48.3	408	7	ADBE62859 Human pro
15	43.5	48.3	415	7	ADBE63989 Rat Prote
16	43.5	48.3	415	7	ADBE63993 Rat Prote
17	43.5	48.3	439	4	ABG08417 Novel hum
18	43.5	48.3	460	3	ABBS6987 Breast an
19	43.5	48.3	460	5	ABP41511 Human ova
20	43	47.8	129	4	AAW84747 Human imm
21	43	47.8	184	6	ABU54886 Metabolic
22	43	47.8	448	2	AAW41516 Cystachio
23	43	47.8	448	2	AAW40920 NHS-5 Pro
24	43	47.8	507	2	AAW42284 Protein w
25	43	47.8	507	5	ABG93112 S. cerevi

26	43	47.8	507	6	ABR53760 Protein s
27	43	47.8	1252	4	AAE02008 Yeast cys
28	43	47.8	1252	6	AAO29518 Yeast cys
29	42.5	47.2	17	6	ABG72113 Viral rep
30	42.5	47.2	18	3	AAV52200 Human la
31	42.5	47.2	18	6	ABG72109 Viral rep
32	42.5	47.2	18	6	ABG72101 Viral rep
33	42.5	47.2	18	6	ABG72114 Viral rep
34	42.5	47.2	19	6	ABG72116 Viral rep
35	42	46.7	935	4	ABB62742 Drosophila
36	41	45.6	980	3	AAW51323 Arabidops
37	41	45.6	1007	2	AAV19957 B. burgdo
38	41	45.6	1031	3	AAW51322 Arabidops
39	41	45.6	1036	2	AAV19956 B. burgdo
40	41	45.6	1036	6	ABU19108 Protein e
41	41	45.6	1277	3	AAW51321 Arabidops
42	40.5	45.0	411	6	ABP70556 Histone d
43	40	44.4	176	4	ABBS64852 Drosophila
44	40	44.4	287	6	ADA36743 Acinetoba
45	40	44.4	379	6	ABU01543 S. pneumo

ALIGNMENTS

RESULT 1	ABG72118 standard; peptide; 18 AA.
ID	ABG72118
AC	ABG72118
DT	28-JAN-2003 (first entry)
XX	Viral replication inhibiting peptide, MOSQUITO.
XX	Viral replication inhibitor; IRES initiated translation; LAP;
XX	internal ribosome entry site initiated translation; La antigen protein;
XX	viral life cycle; antiviral therapy; pharmaceutical; veterinary;
XX	agricultural; horticultural; virucide; mosquito.
XX	Culicoides.
XX	MO200283858-A2.
PD	24-OCT-2002.
XX	12-APR-2002; 2002MO-US011589.
PF	16-APR-2001; 2001US-00836073.
XX	(UYCA-) UNIT CALIFORNIA OFFICE PRESIDENT.
PA	Dasgupta A, Das S, Baidya N;
XX	WPI; 2003-058634/05.
PT	New compound containing acidic and aromatic amino acids, useful as
PT	antiviral therapy in pharmaceutical, veterinary or
PT	agricultural/horticultural applications.
XX	Claim 1; Page 15; 19pp; English.
PS	The present invention relates to peptides and methods of inhibiting the
CC	replication of viruses that utilize internal ribosome entry site (IRES)
CC	initiated translation, and/or inhibiting viruses that utilize the La
CC	antigen protein (LAP) in any phase of their life cycle. The peptides of
CC	the invention compete with LAP and inhibit the utilisation of various
CC	biochemical and physiological functions of LAP required for a productive
CC	life cycle. The methods and compositions are useful as antiviral therapy
CC	in pharmaceutical, veterinary or agricultural/horticultural applications.
CC	ABG72101-ABG72119 represent peptides useful as antiviral agents
XX	Sequence 18 AA;

TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 215:
US-09-402-181B-215

Query Match 46.6%; Score 48.5; DB 4; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.38; 1; Indels 1; Gaps 1;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Oy 7 ICHOIEEYFGD 18
Db 1 ICEQI-EYFGD 11

RESULT 15
US-09-721-456-215
Sequence 215, Application US/09721456
Patent No. 6617110
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Langner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/721,456
FILING DATE: 22-NOV-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION NUMBER: WO PCT/US97/17885

FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 215:
US-09-721-456-215

Query Match 46.6%; Score 48.5; DB 4; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.38; 1; Indels 1; Gaps 1;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Oy 7 ICHOIEEYFGD 18
Db 1 ICEQI-EYFGD 11

Search completed: September 10, 2004, 18:05:10
Job time : 16.771 secs

ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-09-854-050-25

Query Match 46.6%; Score 48.5; DB 3; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.38;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICHOIEEYFGD 18
DB 1 ICEQI-EYFEGD 11

RESULT 13
US-09-430-323-25
Sequence 25, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin H.
Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430.323
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-430-323-25

Query Match 46.6%; Score 48.5; DB 4; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.38;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICHOIEEYFGD 18
DB 1 ICEQI-EYFEGD 11

RESULT 14
US-09-402-181B-215
Sequence 215, Application US/09402181B
Patent No. 6610839
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin H.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 633
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402.181B
FILING DATE: 29-Sep-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ausenius, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US
TELECOMMUNICATION INFORMATION:

TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-25

Query Match 46.6%; Score 48.5; DB 3; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.38;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 7 ICHQIEEYFGD 18
Db 1 ICEQI-EYFGD 11

RESULT 11
US-08-974-549A-215
Sequence 215, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,212
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-215

Query Match 46.6%; Score 48.5; DB 3; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.38;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 7 ICHQIEEYFGD 18
Db 1 ICEQI-EYFGD 11

RESULT 12
US-08-854-050-25
Sequence 25, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6261836el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 536
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536

STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 214:
US-09-402-181B-214

Query Match 55.3%; Score 57.5; DB 4; Length 38;
Best Local Similarity 84.6%; Pred. No. 0.013;
Matches 11; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICHQIEYYFGDF 19
Db 1 ICHQ-XEYFQDF 12

RESULT 9
US-09-721-456-214
Sequence 214, Application US/09721456
Patent No. 6617110

GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.

TITLE OF INVENTION: Human telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/721,456
FILING DATE: 22-NOV-1997
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997

ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 214:
US-09-721-456-214

Query Match 55.3%; Score 57.5; DB 4; Length 38;
Best Local Similarity 84.6%; Pred. No. 0.013;
Matches 11; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICHQIEYYFGDF 19
Db 1 ICHQ-XEYFQDF 12

RESULT 10
US-08-851-843A-25
Sequence 25, Application US/08851843A
Patent No. 6093809

GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.

TITLE OF INVENTION: No. 6093809el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids

RESULT 7
US-08-974-549A-214
Sequence 214, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cecch, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-214
Query Match 55.3%; Score 57.5; DB 3; Length 38;
Best Local Similarity 84.6%; Pred. No. 0.013;
Matches 11; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 7 ICHQIEYYFGDF 19
Db 1 ICHQ-XEYYFGDF 12
RESULT 8
US-09-402-181B-214
Sequence 214, Application US/09402181B
Patent No. 6610839
GENERAL INFORMATION:
APPLICANT: Cecch, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 633
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,181B
FILING DATE: 29-SEP-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ausehuus, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid

PRIOR FILING DATE: 1997-10-06
PRIOR APPLICATION NUMBER: 08/321,427
PRIOR FILING DATE: 1994-10-11
PRIOR APPLICATION NUMBER: 60/086,527
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO: 4
LENGTH: 18
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: B-LAP
OTHER INFORMATION: this peptide is biotinylated
US-09-316-630-4

Query Match 85.1%; Score 88.5; DB 3; Length 18;
Best Local Similarity 94.7%; Pred. No. 5.6e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 AALEKICHQIEYFGDF 19
Db 1 AALEKICHQI-EYFGDF 18

RESULT 3
US-08-475-955-20
Sequence 20, Application US/08475955
Patent No. 6641813
GENERAL INFORMATION:
APPLICANT: Harley, John
TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS OF
NUMBER OF SEQUENCES: 218
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center, 1250 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,955
FILING DATE: June 7, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/667,819
FILING DATE: April 13, 1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/648,205
FILING DATE: January 31, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/472,947
FILING DATE: January 31, 1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMF114CIP(2) DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-873-8795
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Binding-site
LOCATION: 8..15
US-08-475-955-20

Query Match 60.1%; Score 62.5; DB 4; Length 21;
Best Local Similarity 92.3%; Pred. No. 0.0011;
Matches 12; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 7 ICHQIEYFGDF 19
Db 1 ICHQI-EYFGDF 12

RESULT 4
US-08-851-843A-24
Sequence 24, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809e1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-24

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:43:01; Search time 16.771 Seconds
(without alignments)
58.488 Million cell updates/sec

Title: US-09-836-073-16

Perfect score: 104
Sequence: 1 AALEAKICHQIEHYRGDF 19

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/2/iaa/5A COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B COMB.pep:*
3: /cgn2_6/prodata/2/iaa/6A COMB.pep:*
4: /cgn2_6/prodata/2/iaa/6B COMB.pep:*
5: /cgn2_6/prodata/2/iaa/PCTUS COMB.pep:*
6: /cgn2_6/prodata/2/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88.5	85.1	18	3	US-09-316-630-3
2	88.5	85.1	18	3	US-09-316-630-4
3	62.5	60.1	21	4	US-08-475-955-20
4	58	55.8	37	3	US-08-851-843A-24
5	58	55.8	37	3	US-08-854-050-24
6	57.5	55.3	37	4	US-09-430-323-24
7	57.5	55.3	38	4	US-08-974-549A-214
8	57.5	55.3	38	4	US-09-402-181B-214
9	57.5	55.3	38	4	US-09-721-456-214
10	48.5	46.6	38	3	US-08-851-843A-25
11	48.5	46.6	38	3	US-08-974-549A-215
12	48.5	46.6	38	3	US-08-854-050-25
13	48.5	46.6	38	4	US-09-430-323-25
14	48.5	46.6	38	4	US-09-402-181B-215
15	48.5	46.6	38	4	US-09-721-456-215
16	47	45.2	45	1	US-08-205-719-4
17	47	45.2	45	3	US-08-431-517F-6
18	47	45.2	48	2	US-08-431-517F-5
19	47	45.2	48	2	US-08-431-517F-5
20	44	42.3	32	6	5245013-2
21	44	42.3	32	6	5245013-2
22	43	41.3	28	4	US-09-583-681A-4336
23	43	41.3	51	1	US-08-097-829-2
24	42.5	40.9	51	1	US-08-577-803-2
25	42.5	40.9	74	0	US-09-323-872A-23
26	42.5	40.9	74	0	US-09-072-433-15
27	42.5	40.9	86	4	US-09-323-872A-28
			86	4	US-09-072-433-16

28	42	40.4	611	4	US-09-543-681A-5447	Sequence 5447, Ap
29	42	40.4	699	4	US-09-543-681A-5118	Sequence 5118, Ap
30	41.5	39.9	463	4	US-09-543-681A-6745	Sequence 6745, Ap
31	41	39.4	289	4	US-09-543-681A-8235	Sequence 8235, Ap
32	41	39.4	307	4	US-09-328-352-5689	Sequence 5689, Ap
33	41	39.4	410	4	US-09-543-681A-5407	Sequence 5407, Ap
34	41	39.4	497	4	US-09-489-039A-8926	Sequence 8926, Ap
35	40	38.5	41	4	US-09-205-258-970	Sequence 970, App
36	40	38.5	102	4	US-09-198-452A-56	Sequence 56, App1
37	40	38.5	203	4	US-09-328-352-7692	Sequence 7692, Ap
38	40	38.5	214	4	US-09-205-258-965	Sequence 965, App
39	40	38.5	255	4	US-09-107-532A-6662	Sequence 6662, Ap
40	40	38.5	263	2	US-08-864-799-1	Sequence 1, App1
41	40	38.5	1253	1	US-07-920-281C-3	Sequence 3, App1
42	40	38.5	1253	3	US-08-466-277-3	Sequence 3, App1
43	39	37.5	86	4	US-09-300-008B-55	Sequence 55, App1
44	39	37.5	99	4	US-09-543-681A-6663	Sequence 6663, Ap
45	39	37.5	122	4	US-09-732-210-62	Sequence 62, App1

ALIGNMENTS

```

RESULT 1
US-09-316-630-3
Sequence 3, Application US/09316630
Patent No. 6291637
GENERAL INFORMATION:
APPLICANT: Das, Saumitra
APPLICANT: Dasgupta, Asim
TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
FILE REFERENCE: 22000-20548.21
CURRENT APPLICATION NUMBER: US/09/316,630
CURRENT FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: 08/817,953
PRIOR FILING DATE: 1997-10-06
PRIOR APPLICATION NUMBER: 08/321,427
PRIOR FILING DATE: 1994-10-11
PRIOR APPLICATION NUMBER: 60/086,527
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 18
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: IAP
US-09-316-630-3
Query Match 85.1%; Score 88.5; DB 3; Length 18;
Best Local Similarity 94.7%; Pred. No. 5.6e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
DB 1 AALEAKICHQIEHYRGDF 19
1 AALEAKICHQIEHYRGDF 18
RESULT 2
US-09-316-630-4
Sequence 4, Application US/09316630
Patent No. 6291637
GENERAL INFORMATION:
APPLICANT: Das, Saumitra
APPLICANT: Dasgupta, Asim
TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
FILE REFERENCE: 22000-20548.21
CURRENT APPLICATION NUMBER: US/09/316,630
CURRENT FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: 08/817,953

```


DR N-PSDB; AAF22604.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 20; SEQ ID NO 38776; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotide are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp://wipo.int/pub/published_pct_sequences

XX Sequence 439 AA;

SO Query Match 85.1%; Score 88.5; DB 4; Length 439;
 Best Local Similarity 94.7%; Pred. No. 1.4e-05;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
 |||||
 DB 41 AALEAKICHQI-EYYFGDF 58

RESULT 15
 AAB58987 standard; protein; 460 AA.

XX AAB58987;
 XX
 XX 27-MAR-2001 (first entry)
 XX
 DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 695.
 XX
 KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW neoplastic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antidiabetic; antiinflammatory; antilicer; vulnerary; anticonvulsant;
 KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KW cardiovascular disorder; wound healing; neurological disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200055173-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US005881.
 XX
 PR 12-MAR-1999; 99US-0124270P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX

PI Rosen CA, Ruben SM;
 XX WPI; 2000-611515/58.
 DR N-PSDB; AAF21890.

XX New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention, treatment
 PT and diagnosis of cancer, immune disorders, cardiovascular disorders and
 PT neurological diseases.

XX Claim 11; Page 1149-1150; 1299pp; English.

XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive; neurotropic;
 CC neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
 CC antiinflammatory; antilicer; vulnerary; anticonvulsant; antibacterial;
 CC antifungal; antiparasitic and cardiant activity. The polynucleotide and
 CC protein sequences are used in the diagnosis of cancer, particularly
 CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
 CC and agonists may also be used in the diagnosis, prevention and treatment
 CC of immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
 CC cardiovascular disorders such as myocardial ischaemias; wound healing;
 CC neurological diseases such as cerebral anoxia and epilepsy; and
 CC infectious diseases

XX Sequence 460 AA;

SO Query Match 85.1%; Score 88.5; DB 3; Length 460;
 Best Local Similarity 94.7%; Pred. No. 1.4e-05;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
 |||||
 DB 63 AALEAKICHQI-EYYFGDF 80

Search completed: September 10, 2004, 17:51:50
 Job time : 60.5413 secs

CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain; a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain; a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides; a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification; a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 408 AA;

Query Match

Best Local Similarity 85.1%; Score 88.5; DB 7; Length 408;

Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 AALEAKICHQIEYYFGDF 19

Db 11 AALEAKICHQI-EYYFGDF 28

RESULT 11

ADE62859 ID ADE62859 standard; protein; 408 AA.

XX ADE62859;

DT 29-JAN-2004 (first entry)

DE Human Protein P05455; SEQ ID NO 8793.

XX Human; pain; neuronal tissue; gene therapy;

XX spinal segmental nerve injury; chronic constriction injury; CCI;

XX spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI, 2003-268312/26.

XX GENBANK; P05455.

XX New composition comprising two or more isolated polypeptides, useful for

XX preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain; a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain; a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides; a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification; a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 408 AA;

Query Match

Best Local Similarity 85.1%; Score 88.5; DB 7; Length 408;

Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 AALEAKICHQIEYYFGDF 19

Db 11 AALEAKICHQI-EYYFGDF 28

RESULT 12

ADE63989 ID ADE63989 standard; protein; 415 AA.

XX ADE63989;

DT 29-JAN-2004 (first entry)

DE Rat Protein P38656; SEQ ID NO 9935.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI, 2003-268312/26.

XX GENBANK; P38656.

XX New composition comprising two or more isolated polypeptides, useful for

XX preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 408 AA;

Query Match 85.1%; Score 88.5; DB 7; Length 408;
Best Local Similarity 94.7%; Pred. No. 1.3e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
DB 11 AALEAKICHQI-EYYFGDF 28

RESULT 9
ADD46272
ID ADD46272 standard; protein; 408 AA.
XX
XX ADD46272;

XX 29-JAN-2004 (first entry)

DE Human Protein P05455, SEQ ID NO 11947.

XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX MO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; P05455.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; P05455.

CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 408 AA;

Query Match 85.1%; Score 88.5; DB 7; Length 408;
Best Local Similarity 94.7%; Pred. No. 1.3e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
DB 11 AALEAKICHQI-EYYFGDF 28

RESULT 10
ADE63995
ID ADE63995 standard; protein; 408 AA.
XX
XX ADE63995;

XX 29-JAN-2004 (first entry)

DE Human Protein P05455, SEQ ID NO 9941.

XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX MO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; P05455.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 408 AA;

Query Match 85.1%; Score 88.5; DB 7; Length 408;
Best Local Similarity 94.7%; Pred. No. 1.3e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
DB 11 AALEAKICHQI-EYYFGDF 28

RESULT 10
ADE63995
ID ADE63995 standard; protein; 408 AA.
XX
XX ADE63995;

XX 29-JAN-2004 (first entry)

DE Human Protein P05455, SEQ ID NO 9941.

XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX MO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; P05455.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

is clinically related to the Ro protein that is highly common among autoimmune patients. La and Ro antigens sometimes reside on the same cellular ribonucleoprotein particle; most La patients contain some Ro antibodies and vice versa. La cDNA has been isolated from a human liver library. (Updated on 25-MAR-2003 to correct PF field.)

Sequence 408 AA;

Query Match 85.1%; Score 88.5; DB 2; Length 408;
Best Local Similarity 94.7%; Pred. No. 1.3e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

1 AALEKICHQIEEYFGDF 19
11 AALEKICHQI-EYFGDF 28

RESULT 7
ABP65252

ID ABP65252 standard; protein; 408 AA.

AC ABP65252;

DT 12-NOV-2002 (first entry)

DE Hypoxia-regulated protein #126.

Cycostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;
antiinflammatory; vulnery; gynecological; ophthalmological; vaccine;
hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;
ischaemic condition; reperfusion injury; retinopathy; neonatal stress;
preclapmia; atherosclerosis; inflammatory condition; wound healing;
inflammation; erythropoiesis; hair loss; human.

Homo sapiens.

WO200246465-A2.

13-JUN-2002.

10-DEC-2001; 2001WO-GB005458.

08-DEC-2000; 2000GB-00030076.

08-FEB-2001; 2001GB-00003156.

25-OCT-2001; 2001GB-00025666.

(OXFO-) OXFORD BIOMEDICA UK LTD.

White J, Mundy CR, Ward NR, Krige D, Kingsman SM, Harris RA;

Rayner WN;

Identifying a gene involved in disease for treating hypoxia-regulated

conditions, comprises comparing the transcriptome/proteome of two cell

types under different conditions and identifying a differentially

regulated gene.

Claim 35; Page 425; 538pp; English.

The present invention relates to methods for identifying genes and
proteins that are implicated in a specific disease or physiological
condition. The method comprises comparing the transcriptome/proteome of a
specialised cell type implicated in a disease or condition with that of a
second specialised cell type, under two experimental conditions, and
identifying a gene that is differentially regulated in the two
specialised cell types under experimental conditions. ABV7873-ABV78116
and ABP65061-ABP65257 were identified using the methods of the invention.
The coding sequences and proteins are useful for treating a disease in a
patient, for manufacture of a medicament for treating hypoxia-regulated
conditions, and for regulating tumorigenesis, angiogenesis, apoptosis,
biological response to hypoxia conditions, or hypoxia-associated
pathology in a patient. The coding sequences and proteins are also useful

for monitoring the therapeutic treatment of a disease or physiological
condition, such as cancer, ischaemic conditions, reperfusion injury,
retinopathy, neonatal stress, preclapmia, atherosclerosis, inflammatory
conditions, wound healing, inflammation, erythropoiesis or hair loss

Sequence 408 AA;

Query Match 85.1%; Score 88.5; DB 5; Length 408;
Best Local Similarity 94.7%; Pred. No. 1.3e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

1 AALEKICHQIEEYFGDF 19
11 AALEKICHQI-EYFGDF 28

RESULT 8
ADE63991

ID ADE63991 standard; protein; 408 AA.

AC ADE63991;

DT 29-JAN-2004 (first entry)

DE Human Protein P05455, SEQ ID NO 9937.

Human; pain; neuronal tissue; gene therapy;
spinal segmental nerve injury; chronic constriction injury; CCI;
spared nerve injury; SNI; Chung.

Homo sapiens.

WO2003016475-A2.

27-FEB-2003.

14-AUG-2002; 2002WO-US025765.

14-AUG-2001; 2001US-0312147P.

01-NOV-2001; 2001US-0346382P.

26-NOV-2001; 2001US-0333347P.

(GEHO) GEN HOSPITAL CORP.

(PARB) BAYER AG.

Woolf C, D'urso D, Befort K, Costigan M;

GENBANK; P05455.

New composition comprising two or more isolated polypeptides, useful for

preparing a medicament for treating pain in an animal.

Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat
or human polynucleotides or a polynucleotide which represents a fragment,
derivative or allelic variation of the nucleic acid sequence. Also
claimed are a vector comprising the novel polynucleotide, a host cell
comprising the vector, a method for identifying a nucleotide sequence
which is differentially regulated in an animal subjected to pain and a
kit to perform the method, an array, a method for identifying an agent
that increases or decreases the expression of the polynucleotide sequence
that is differentially expressed in neuronal tissue of a first animal
subjected to pain, a method for identifying a compound which regulates
the expression of a polynucleotide sequence which is differentially
expressed in an animal subjected to pain, a method for identifying a
compound that regulates the activity of one or more of the
polynucleotides, a method for producing a pharmaceutical composition, a
method for identifying a compound or small molecule that regulates the
activity for identifying a compound or more of the polypeptides given in the
specification, a method for identifying a compound useful in treating
pain and a pharmaceutical composition comprising the one or more

KW internal ribosome entry site initiated translation; La antigen protein;
 KW viral life cycle; antiviral therapy; pharmaceutical; veterinary;
 KM agricultural; horticultural; virucide; bovine.
 XX Bovinae.
 OS
 XX WO200283858-A2.
 PN
 XX 24-OCT-2002.
 PD
 XX 12-APR-2002; 2002WO-US011589.
 PF
 XX 16-APR-2001; 2001US-00836073.
 PR
 XX (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.
 PA
 XX Dasgupta A, Das S, Balda N;
 PI
 XX WPI; 2003-058634/05.
 DR
 XX New compound containing acidic and aromatic amino acids, useful as
 PT antiviral therapy in pharmaceutical, veterinary or
 PT agricultural/horticultural applications.
 PS
 XX Disclosure; Page 6; 19pp; English.
 XX
 CC The present invention relates to peptides and methods of inhibiting the
 CC replication of viruses that utilize internal ribosome entry site (IRES)
 CC initiated translation, and/or inhibiting viruses that utilize the La
 CC antigen protein (IAP) in any phase of their life cycle. The peptides of
 CC the invention compete with IAP and inhibit the utilisation of various
 CC biochemical and physiological functions of IAP required for a productive
 CC life cycle. The methods and compositions are useful as antiviral therapy
 CC in pharmaceutical, veterinary or agricultural/horticultural applications.
 CC ABG72101-ABG72119 represent peptides useful as antiviral agents
 XX
 SQ Sequence 18 AA;
 XX
 QY Query Match 85.1%; Score 88.5; DB 6; Length 18;
 Best Local Similarity 94.7%; Pred. No. 4.3e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 XX
 DB 1 AALEAKICHOIEHYRGDF 19
 1 AALEAKICHOI-EYRGDF 18
 XX
 RESULT 5
 AAG01351
 ID AAG01351 standard; protein; 92 AA.
 AC
 XX AAG01351;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 5432.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 OS
 XX EP1033401-A2.
 PN
 XX 06-SEP-2000.
 PD
 XX 21-FEB-2000; 2000EP-00200610.
 PF
 XX 26-FEB-1999; 99US-0122487P.
 PR
 XX (GEST) GENSET.
 PA
 XX Dumas Malne Edwards J, Duclert A, Giordano J;
 PI

XX
 DR WPI; 2000-500381/45.
 DR N-PSDB; AAC01357.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
 XX
 PS Claim 13; SEQ ID NO 5432; 71pp + Sequence Listing; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number of
 CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
 CC prepared from total human RNAs or polyA+ RNAs derived from 30 different
 CC tissues. EST sequences usually correspond mainly to the 3' untranslated
 CC region (UTR) of the mRNA because they are often obtained from oligo-dT
 CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
 CC sequences derived from the 5' ends of mRNAs and even in those cases where
 CC longer cDNA sequences have been obtained, the full 5' UTR is rarely
 CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can
 CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
 CC are also used in diagnostic, forensic, gene therapy and chromosome
 CC mapping procedures. They are used to obtain upstream regulatory sequences
 CC and to design expression and secretion vectors
 XX
 SQ Sequence 92 AA;
 XX
 QY Query Match 85.1%; Score 88.5; DB 3; Length 92;
 Best Local Similarity 94.7%; Pred. No. 2.5e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 XX
 DB 1 AALEAKICHOIEHYRGDF 19
 11 AALEAKICHOI-EYRGDF 28
 XX
 RESULT 6
 AAM03716
 ID AAM03716 standard; protein; 408 AA.
 AC
 XX AAM03716;
 XX
 DT 25-MAR-2003 (revised)
 DT 12-MAR-1997 (first entry)
 XX
 DE Human autoantigen Ia(SS-B).
 XX
 KW Autoimmune disease; Ia autoantigen; Sjogren's syndrome;
 KW systemic lupus erythematosus; diagnosis.
 XX
 OS Homo sapiens.
 OS
 XX US5541291-A.
 PN
 XX 30-JUL-1996.
 PD
 XX 27-MAY-1987; 87US-00054871.
 PF
 XX 31-DEC-1984; 84US-00687908.
 PR
 XX (UYDU-) UNIV DUKE.
 PA
 XX Keene JD;
 PI
 XX WPI; 1996-362015/36.
 DR
 XX Auto-antigen U2-RNP, associated with severe polymyositis-scleroderma
 PT overlap syndrome - useful for diagnosis and treatment of autoimmune
 PT diseases.
 XX
 PS Disclosure; Col 15-16; 21pp; English.
 XX
 CC The human lupus antigen (Ia) is diagnostic for Sjogren's syndrome, as
 CC well as occurring in systemic lupus erythematosus patients. The Ia protein

Query Match 100.0%; Score 104; DB 6; Length 19;
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AALEAKICHQIEEYFGDF 19
 |||||
 1 AALEAKICHQIEEYFGDF 19

RESULT 2
 AAY52200
 ID AAY52200 standard; peptide; 18 AA.
 XX
 AC AAY52200;
 XX
 DT 14-MAR-2000 (first entry)
 XX
 DE Human Ia autoantigen peptide (LAP).
 XX

La autoantigen; LAP, internal ribosome entry site; IRES; translation;
 viral replication; IRNA; antiviral agent; picornavirus; flavivirus;
 coronavirus; hepatitis virus; adenovirus; coxsackie virus;
 parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
 foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
 vesicular stomatitis virus.

OS Homo sapiens.
 XX
 XX MO9961613-A2.
 XX
 PD 02-DEC-1999.
 XX
 XX 21-MAY-1999; 99WO-US011281.
 XX
 PR 22-MAY-1998; 98US-0086527P.
 XX
 PA (REGC). UNIV CALIFORNIA.
 XX

PI Das S, Dasgupta A;
 XX
 DR WPI; 2000-062712/05.
 XX

PT New yeast inhibitory peptide useful for inhibiting viral protein
 translation and replication.
 XX

PS Claim 5; Page 57; 81pp; English.
 XX

CC This sequence is the Ia autoantigen binding domain (LAP). LAP is a
 cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors from
 CC binding at the site. The peptide can be used alone or in combination with
 CC an inhibitor RNA (IRNA see AAY5200). The LAP peptide is useful as an
 CC antiviral agent, which works through the inhibition of mRNA translation,
 CC especially viral mRNA. Examples of viruses which can be inhibited are
 CC picornavirus, flavivirus, coronavirus, hepatitis A B or C viruses,
 CC rhabdovirus, adenovirus, and parainfluenza virus, poliovirus, rhinovirus,
 CC coxsackie virus, encephalomyocarditis virus, foot-and-mouth disease
 CC virus, echo virus, hepatitis C virus, infectious bronchitis virus, duck
 CC and human hepatitis B virus, and vesicular stomatitis virus. The peptide
 CC also inhibits replication of the above viruses. The LAP peptide
 CC selectively inhibits viral protein translation, and is therefore not
 CC toxic to the host cell
 CC
 XX

XX Sequence 18 AA;
 SQ

Query Match 85.1%; Score 88.5; DB 3; Length 18;
 Best Local Similarity 94.7%; Pred. No. 4.3e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 AALEAKICHQIEEYFGDF 19

DB 1 AALEAKICHQI-EYFGDF 18
 |||||
 1 AALEAKICHQI-EYFGDF 18

RESULT 3
 ABG72101
 ID ABG72101 standard; peptide; 18 AA.
 XX
 AC ABG72101;
 XX

XX 28-JAN-2003 (first entry)
 XX
 DE Viral replication inhibiting peptide, LAP.
 XX

Viral replication inhibitor; IRES initiated translation; LAP;
 internal ribosome entry site initiated translation; Ia antigen protein;
 viral life cycle; antiviral therapy; pharmaceutical; veterinary;
 agricultural; horticultural; virucide.

OS Unidentified.
 XX
 XX WO200283858-A2.
 XX
 PD 24-OCT-2002.
 XX
 PR 12-APR-2002; 2002WO-US011589.
 XX

PR 16-APR-2001; 2001US-00836073.
 XX

PA (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.
 XX

PI Dasgupta A, Das S, Balady N;
 XX

DR WPI; 2003-058634/05.
 XX

PT New compound containing acidic and aromatic amino acids, useful as
 antiviral therapy in pharmaceutical, veterinary or
 PT agricultural/horticultural applications.
 XX

PS Claim 1; Page 15; 19pp; English.
 XX

CC The present invention relates to peptides and methods of inhibiting the
 CC replication of viruses that utilise internal ribosome entry site (IRES)
 CC initiated translation, and/or inhibiting viruses that utilise the Ia
 CC antigen protein (LAP) in any phase of their life cycle. The peptides of
 CC the invention compete with LAP and inhibit the utilisation of various
 CC biochemical and physiological functions of LAP required for a productive
 CC life cycle. The methods and compositions are useful as antiviral therapy
 CC in pharmaceutical, veterinary or agricultural/horticultural applications.
 CC ABG72101-ABG72119 represent peptide useful as antiviral agents
 XX

XX Sequence 18 AA;
 SQ

Query Match 85.1%; Score 88.5; DB 6; Length 18;
 Best Local Similarity 94.7%; Pred. No. 4.3e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 AALEAKICHQIEEYFGDF 19
 |||||
 1 AALEAKICHQI-EYFGDF 18

DB 1 AALEAKICHQI-EYFGDF 18
 |||||
 1 AALEAKICHQI-EYFGDF 18

XX Sequence 18 AA;
 XX

XX 28-JAN-2003 (first entry)
 XX
 DT Viral replication inhibiting peptide, BOVINE.
 XX
 DE Viral replication inhibitor; IRES initiated translation; LAP;
 XX

XX Viral replication inhibitor; IRES initiated translation; LAP;
 XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using: sw model

Run on: September 10, 2004, 16:58:20 ; Search time 59.4413 Seconds
(without alignments)
90.314 Million cell updates/sec

Title: US-09-836-073-16

Perfect score: 104

Sequence: 1 AALEKICHOIEEYFGDP 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20048:*
8: geneseqp20058:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	19	ABG72116	Abg72116 Viral rep
2	88.5	85.1	18	AAAY52200	Hay52200 Human Ia
3	88.5	85.1	18	ABG72101	Abg72101 Viral rep
4	88.5	85.1	18	ABG72114	Abg72114 Viral rep
5	88.5	85.1	92	AAAG01351	AAg01351 Human sec
6	88.5	85.1	408	AAAG03716	AAg03716 Human aut
7	88.5	85.1	408	ABP65252	ABp65252 Hypoxia-r
8	88.5	85.1	408	ADDE63991	ADde63991 Human pro
9	88.5	85.1	408	ADDA46272	ADda46272 Human pro
10	88.5	85.1	408	ADDE63995	ADde63995 Human pro
11	88.5	85.1	408	ADDE63989	ADde63989 Human pro
12	88.5	85.1	415	ADDE63989	ADde63989 Rat Prote
13	88.5	85.1	415	ADDE63993	ADde63993 Rat Prote
14	88.5	85.1	439	ABG08417	ABg08417 Novel hum
15	88.5	85.1	460	AAAB58987	AAb58987 Breast an
16	88.5	85.1	460	ABPA41511	ABp41511 Human ova
17	84.5	81.2	17	ABG72113	Abg72113 Viral rep
18	83.5	80.3	18	ABG72111	Abg72111 Viral rep
19	82.5	79.3	18	ABG72109	Abg72109 Viral rep
20	80.5	77.4	18	ABG72112	Abg72112 Viral rep
21	79.5	76.4	18	ABG72110	Abg72110 Viral rep
22	76.5	73.6	18	ABG72102	Abg72102 Viral rep
23	76.5	73.6	18	ABG72103	Abg72103 Viral rep
24	76.5	73.6	18	ABG72104	Abg72104 Viral rep
25	75.5	72.6	18	ABG72108	Abg72108 Viral rep

ALIGNMENTS

26	73.5	70.7	18	6	ABG72107	Abg72107 Viral rep
27	65.5	63.0	18	6	ABG72115	Abg72115 Viral rep
28	64.5	62.0	18	6	ABG72105	Abg72105 Viral rep
29	63	60.6	18	6	ABG72106	Abg72106 Viral rep
30	62.5	60.1	21	2	AAK43394	AAk43394 La/SSB ep
31	49	47.1	2169	4	AAAG91701	AAg91701 C glutamL
32	47	45.2	456	2	AAW41659	AAw41659 Rabbit LP
33	47	45.2	482	2	AAW41659	AAw41659 Lapine gr
34	47	45.2	482	2	AAW53461	AAw53461 Lapine po
35	47	45.2	482	2	AAW40813	AAw40813 Rabbit 11
36	47	45.2	482	2	AAW78355	AAw78355 Rabbit 11
37	46.5	44.7	16	6	ABG72119	ABg72119 Viral rep
38	46	44.2	352	6	ABR53809	ABr53809 Protein B
39	45.5	43.8	132	3	AAAB51905	AAa51905 Human sec
40	45.5	43.8	425	4	AAW25631	AAw25631 Human pro
41	45.5	43.8	643	4	AAAB99540	AAa99540 Prelimina
42	45.5	43.8	643	7	ADC21971	ADc21971 Human cad
43	45.5	43.8	675	6	ABR43633	ABr43633 Mouse CLA
44	45.5	43.8	863	7	ADC31112	ADc31112 Human nov
45	45.5	43.8	2047	4	AAAB99541	AAa99541 Human CLA

RESULT 1
ID ABG72116 standard; peptide; 19 AA.
AC ABG72116;
XX
XX 28-JUN-2003 (first entry)
XX
XX
XX Viral replication inhibiting peptide, RAT.
DE
XX Viral replication inhibitor; IRBS initiated translation; LAP;
KW internal ribosome entry site initiated translation; Ia antigen protein;
KW viral life cycle; antiviral therapy; pharmaceutical; veterinary;
KW agricultural; horticultural; virucide; rat.
XX
XX Rattus sp.
OS
XX
XX WO200283858-A2.
XX
XX 24-OCT-2002.
XX
XX 12-APR-2002; 2002WO-US011589.
XX
XX 16-APR-2001; 2001US-00836073.
XX
XX (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.
XX
XX Dasgupta A, Das S, Balda N;
PI
XX
XX WPI; 2003-058634/05.
XX
XX New compound containing acidic and aromatic amino acids, useful as
PT antiviral therapy in pharmaceutical, veterinary or
PT agricultural/horticultural applications.
XX
XX Claim 1; Page 15; 19pp; English.
XX
XX The present invention relates to peptides and methods of inhibiting the
XX replication of viruses that utilize internal ribosome entry site (IRES)
XX initiated translation, and/or inhibiting viruses that utilize the Ia
XX antigen protein (IAP) in any phase of their life cycle. The peptides of
XX the invention compete with IAP and inhibit the utilisation of a productive
XX biochemical and physiological functions of IAP required for a virus
XX life cycle. The methods and compositions are useful as antiviral therapy
XX in pharmaceutical, veterinary or agricultural/horticultural applications.
XX ABG72101-ABG72119 represent peptides useful as antiviral agents
XX
XX Sequence 19 AA;
SQ

CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 415 AA;

Query Match 76.0%; Score 76; DB 7; Length 415;

Best Local Similarity 81.2%; Pred. No. 0.0022; Mismatches 2; Indels 0; Gaps 0;

Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 3 LDTKICEQIEYFGDF 18
 | : ||| |||||
 Db 13 LEAKICHQIEYFGDF 28

RESULT 15

ABG08417 ABG08417 standard; protein; 439 AA.

AC ABG08417;

DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #8408.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

XX PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

XX WIPI; 2001-639362/73.

DR N-PSDB; AAS72604.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

PS Claim 20; SEQ ID NO 38776; 103bp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC patent did not appear in the printed specification. Note: The sequence data for this
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 439 AA;

Query Match 76.0%; Score 76; DB 4; Length 439;

Best Local Similarity 81.2%; Pred. No. 0.0023; Mismatches 2; Indels 0; Gaps 0;

Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 3 LDTKICEQIEYFGDF 18
 | : ||| |||||
 Db 43 LEAKICHQIEYFGDF 58

Search completed: September 10, 2004, 17:51:49
 Job time : 57.4128 secs

polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 408 AA;

Query Match 76.0%; Score 76; DB 7; Length 408;
Best Local Similarity 81.2%; Pred. No. 0.0022;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYFGDF 18
|:|||||
Db 13 LEAKICHQIEYFGDF 28

RESULT 13
ADE63989
ID ADE63989 standard; protein; 415 AA.
XX ADE63989;

29-JAN-2004 (first entry)

Rat Protein P38656, SEQ ID NO 9935.

Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

Rattus norvegicus.

MO2003016475-A2.

27-FEB-2003.

14-AUG-2002; 2002WO-US025765.

14-AUG-2001; 2001US-0312147P.

01-NOV-2001; 2001US-0346382P.

26-NOV-2001; 2001US-033347P.

(GEHO) GEN HOSPITAL CORP.

(FARB) BAYER AG.

Woolf C, D'urso D, Befort K, Costigan M;

WPI; 2003-268312/26.

GENBANK; P38656.

Claim 1; Page; 1017pp; English.

New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a

method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a rat protein (shown in table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 415 AA;

Query Match 76.0%; Score 76; DB 7; Length 415;
Best Local Similarity 81.2%; Pred. No. 0.0022;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYFGDF 18
|:|||||
Db 13 LEAKICHQIEYFGDF 28

RESULT 14
ADE63993
ID ADE63993 standard; protein; 415 AA.
XX ADE63993;

29-JAN-2004 (first entry)

Rat Protein P38656, SEQ ID NO 9939.

Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

Rattus norvegicus.

MO2003016475-A2.

27-FEB-2003.

14-AUG-2002; 2002WO-US025765.

14-AUG-2001; 2001US-0312147P.

01-NOV-2001; 2001US-0346382P.

26-NOV-2001; 2001US-033347P.

(GEHO) GEN HOSPITAL CORP.

(FARB) BAYER AG.

Woolf C, D'urso D, Befort K, Costigan M;

WPI; 2003-268312/26.

GENBANK; P38656.

Claim 1; Page; 1017pp; English.

New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates

SQL Sequence 408 AA;

Query Match 76.0%; Score 76; DB 7; Length 408;
Best Local Similarity 81.2%; Pred. No. 0.0022;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYYPGDF 18
|:|||||
DB 13 LEAKICHQIEYYPGDF 28

RESULT 11
ADE63995
ID ADE63995 standard; protein; 408 AA.
XX
AC ADE63995;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein P05455, SEQ ID NO 9941.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GENO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR MPI; 2003-268312/26.
DR GENBANK; P05455.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of

CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

SQL Sequence 408 AA;

Query Match 76.0%; Score 76; DB 7; Length 408;
Best Local Similarity 81.2%; Pred. No. 0.0022;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYYPGDF 18
|:|||||
DB 13 LEAKICHQIEYYPGDF 28

RESULT 12
ADE62859
ID ADE62859 standard; protein; 408 AA.
XX
AC ADE62859;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein P05455, SEQ ID NO 8793.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GENO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR MPI; 2003-268312/26.
DR GENBANK; P05455.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more

RESULT 9
 ADE63991
 ID ADE63991 standard; protein; 408 AA.
 XX
 AC ADE63991;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human Protein P05455, SEQ ID NO 9937.
 XX
 DE Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX
 OS Homo sapiens.
 XX
 PN WO2003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 14-AUG-2002; 2002WO-US025765.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PA (GENO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 PI Woolf C, D'urso D, Befort K, Costigan M;
 DR WPI; 2003-268312/26.
 DR GENBANK; P05455.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 PS Claim 1; Page; 1017pp; English.
 XX
 CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 408 AA;
 XX
 Query Match 76.0%; Score 76; DB 7; Length 408;
 Best Local Similarity 81.2%; Pred. NO. 0.0022;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LPTKICEQLEYFGDF 18
 DB 13 LEAKICHOLEYFGDF 28
 RESULT 10
 ADD46272
 ID ADD46272 standard; protein; 408 AA.
 XX
 AC ADD46272;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human Protein P05455, SEQ ID NO 11947.
 XX
 DE Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX
 OS Homo sapiens.
 XX
 PN WO2003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 14-AUG-2002; 2002WO-US025765.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PA (GENO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 PI Woolf C, D'urso D, Befort K, Costigan M;
 DR WPI; 2003-268312/26.
 DR GENBANK; P05455.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 PS Claim 1; Page; 1017pp; English.
 XX
 CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX

CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
CC sequences derived from the 5' ends of mRNAs and even in those cases where
CC longer cDNA sequences have been obtained, the full 5' UTR is rarely
CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can
CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
CC are also used in diagnostic, forensic, gene therapy and chromosome
CC mapping procedures. They are used to obtain upstream regulatory sequences
CC and to design expression and secretion vectors

XX Sequence 92 AA;

Query Match 76.0%; Score 76; DB 3; Length 92;

Best Local Similarity 81.2%; Pred. No. 0.00042; Mismatches 2; Indels 0; Gaps 0;

Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 3 LDTKICEQIEYFGDF 18
13 LEAKICHOIEYFGDF 28

RESULT 7
AAW03716 standard; protein; 408 AA.
XX AAW03716;
XX AAW03716;
XX 25-MAR-2003 (revised)
DT 12-MAR-1997 (first entry)
XX 12-MAR-1997 (first entry)
DE Human autoantigen Ia (SS-B).

XX Autoimmune disease; La autoantigen; Sjogren's syndrome;
KM systemic lupus erythematosus; diagnosis.

XX Homo sapiens.

OS US5541291-A.

PN 30-JUL-1996.

PD 27-MAY-1987; 87US-00054871.

PF 31-DEC-1984; 84US-00687908.

PR (UYDU-) UNITV DUKE.

PA Keene JD;

PI WPI; 1996-362015/36.

PT Auto-antigen U2-RNP, associated with severe polymyositis-scleroderma
PT overlap syndrome - useful for diagnosis and treatment of autoimmune
PT diseases.

PS Disclosure; Col 15-16; 21pp; English.

XX The human lupus antigen (La) is diagnostic for Sjogren's syndrome, as
CC well as occurring in systemic lupus erythematosus patients. The La protein
CC is clinically related to the Ro protein that is highly common among
CC autoimmune patients. La and Ro antigens sometimes reside on the same
CC cellular ribonucleoprotein particle; most La patients contain some Ro
CC antibodies and vice versa. La cDNA has been isolated from a human liver
CC library. (Updated on 25-MAR-2003 to correct PF field.)

XX Sequence 408 AA;

Query Match 76.0%; Score 76; DB 2; Length 408;

Best Local Similarity 81.2%; Pred. No. 0.0022; Mismatches 2; Indels 0; Gaps 0;

Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 3 LDTKICEQIEYFGDF 18
13 LEAKICHOIEYFGDF 28

RESULT 8
ABP65252 standard; protein; 408 AA.

XX ABP65252;

DT 12-NOV-2002 (first entry)

DE Hypoxia-regulated protein #126.

XX Cytoprotective; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;
KM antiinflammatory; vulnerrary; gynecological; ophthalmological; vaccine;
KM hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;
KM ischaemic condition; reperfusion injury; retinopathy; neonatal stress;
KM preclapamsia; atherosclerosis; inflammatory condition; wound healing;
KM inflammation; erythropoiesis; hair loss; human.

XX Homo sapiens.

PN WO200246465-A2.

PD 13-JUN-2002.

PF 10-DEC-2001; 2001WO-GB005458.

PR 08-DEC-2000; 2000GB-00030076.

PR 08-FEB-2001; 2001GB-00003156.

PR 25-OCT-2001; 2001GB-00025666.

PA (OXFO-) OXFORD BIOMEDICA UK LTD.

XX White J, Mundy CR, Ward NR, Krige D, Kingman SM, Harris RA;

PI Rayner WN;

DR WPI; 2002-627238/67.

PT Identifying a gene involved in disease for treating hypoxia-regulated

PT conditions, comprises comparing the transcriptome/proteome of two cell

PT types under different conditions and identifying a differentially

PT regulated gene.

PS Claim 35; Page 425; 538pp; English.

XX The present invention relates to methods for identifying genes and

XX proteins that are implicated in a specific disease or physiological

XX condition. The method comprises comparing the transcriptome/proteome of a

XX specialised cell type implicated in a disease or condition with that of a

XX second specialised cell type, under two experimental conditions, and

XX identifying a gene that is differentially regulated in the two

XX specialised cell types under experimental conditions. ABV7873-ABV78116

XX and ABP65061-ABP65257 were identified using the methods of the invention.

XX The coding sequences and proteins are useful for treating a disease in a

XX patient, for manufacture of a medicament for treating hypoxia-regulated

XX conditions, and for regulating tumorigenesis, angiogenesis, apoptosis,

XX biological response to hypoxia conditions, or hypoxic-associated

XX pathology in a patient. The coding sequences and proteins are also useful

XX for monitoring the therapeutic treatment of a disease or physiological

XX condition, such as cancer, ischaemic conditions, reperfusion injury,

XX retinopathy, neonatal stress, preclapamsia, atherosclerosis, inflammatory

XX conditions, wound healing, inflammation, erythropoiesis or hair loss

XX Sequence 408 AA;

Query Match 76.0%; Score 76; DB 5; Length 408;

Best Local Similarity 81.2%; Pred. No. 0.0022; Mismatches 2; Indels 0; Gaps 0;

Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 3 LDTKICEQIEYFGDF 18
13 LEAKICHOIEYFGDF 28

KW internal ribosome entry site initiated translation; La antigen protein;
 KW viral life cycle; antiviral therapy; pharmaceutical; veterinary;
 KW agricultural; horticultural; virucide.
 XX Unidentified.
 XX OS
 XX WO200283858-A2.
 XX PN
 XX 24-OCT-2002.
 XX PD
 XX 12-APR-2002; 2002WO-US011589.
 XX PF
 XX 16-APR-2001; 2001US-00836073.
 XX PR
 XX (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.
 XX PA
 XX Dasgupta A, Das S, Baidya N;
 XX PI
 XX WPI; 2003-058634/05.
 XX DR
 XX New compound containing acidic and aromatic amino acids, useful as
 XX PT antiviral therapy in pharmaceutical, veterinary or
 XX PR agricultural/horticultural applications.
 XX PS
 XX Claim 1; Page 15; 19pp; English.
 XX CC The present invention relates to peptides and methods of inhibiting the
 XX CC replication of viruses that utilize internal ribosome entry site (IRES)
 XX CC initiated translation, and/or inhibiting viruses that utilize the La
 XX CC antigen protein (LAP) in any phase of their life cycle. The peptides of
 XX CC the invention compete with LAP and inhibit the utilisation of various
 XX CC biochemical and physiological functions of LAP required for a productive
 XX CC life cycle. The methods and compositions are useful as antiviral therapy
 XX CC in pharmaceutical, veterinary or agricultural/horticultural applications.
 XX CC ABG72101-ABG72119 represent peptides useful as antiviral agents
 XX
 SO Sequence 18 AA:
 Query Match 76.0%; Score 76; DB 6; Length 18;
 Best Local Similarity 81.2%; Pred. No. 7e-05;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 3 LDTKICEQTEYVFGDF 18
 DB 3 LEAKICHQIEYVFGDF 18
 DB
 RESULT 5
 ABG72114
 ID ABG72114 standard; peptide; 18 AA.
 XX AC
 XX ABG72114;
 XX
 DT 28-JAN-2003 (first entry)
 XX
 DE Viral replication inhibiting peptide, BOVINE.
 XX
 XX Viral replication inhibitor; IRES initiated translation; LAP;
 KW internal ribosome entry site initiated translation; La antigen protein;
 KW viral life cycle; antiviral therapy; pharmaceutical; veterinary;
 KW agricultural; horticultural; virucide; bovine.
 XX
 XX Bovinae.
 XX OS
 XX WO200283858-A2.
 XX PN
 XX 24-OCT-2002.
 XX PD
 XX 12-APR-2002; 2002WO-US011589.
 XX PF
 XX 16-APR-2001; 2001US-00836073.
 XX PR
 XX (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.
 XX PA

XX
 XX Dasgupta A, Das S, Baidya N;
 XX PI
 XX WPI; 2003-058634/05.
 XX DR
 XX New compound containing acidic and aromatic amino acids, useful as
 XX PT antiviral therapy in pharmaceutical, veterinary or
 XX PR agricultural/horticultural applications.
 XX PS
 XX Disclosure; Page 6; 19pp; English.
 XX CC The present invention relates to peptides and methods of inhibiting the
 XX CC replication of viruses that utilize internal ribosome entry site (IRES)
 XX CC initiated translation, and/or inhibiting viruses that utilize the La
 XX CC antigen protein (LAP) in any phase of their life cycle. The peptides of
 XX CC the invention compete with LAP and inhibit the utilisation of various
 XX CC biochemical and physiological functions of LAP required for a productive
 XX CC life cycle. The methods and compositions are useful as antiviral therapy
 XX CC in pharmaceutical, veterinary or agricultural/horticultural applications.
 XX CC ABG72101-ABG72119 represent peptides useful as antiviral agents
 XX
 SO Sequence 18 AA:
 Query Match 76.0%; Score 76; DB 6; Length 18;
 Best Local Similarity 81.2%; Pred. No. 7e-05;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 3 LDTKICEQTEYVFGDF 18
 DB 3 LEAKICHQIEYVFGDF 18
 DB
 RESULT 6
 AAG01351
 ID AAG01351 standard; protein; 92 AA.
 XX AC
 XX AAG01351;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 5432.
 XX
 XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 KW
 OS Homo sapiens.
 XX
 XX EP1033401-A2.
 XX PN
 XX 06-SEP-2000.
 XX PD
 XX 21-FEB-2000; 2000EP-00200610.
 XX PF
 XX 26-FEB-1999; 99US-0122487P.
 XX PR
 XX (GEST) GENSET.
 XX PA
 XX Dumas Malne Edwards J, Duclert A, Giordano J;
 XX PI
 XX WPI; 2000-500381/45.
 XX DR
 XX N-PSDB; AAC01357.
 XX
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 XX PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
 XX PS
 XX Claim 13; SEQ ID NO 5432; 71pp + Sequence Listing; English.
 XX CC The present sequence is a polypeptide encoded by one of a large number of
 XX CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
 XX CC prepared from total human RNAs or polyA+ RNAs derived from 30 different
 XX CC tissues. EST sequences usually correspond mainly to the 3' untranslated
 XX CC region (UTR) of the mRNA because they are often obtained from oligo-dT

Query Match 100.0%; Score 100; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 1.3e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LDLPKICEQIEYFGDF 18
 |||||
 DB 1 LDLPKICEQIEYFGDF 18

RESULT 2
 ABG72113
 ID ABG72113 standard; peptide; 17 AA.
 XX
 AC ABG72113;
 XX
 DT 28-JAN-2003 (first entry)
 XX
 DE Viral replication inhibiting peptide, MOUSE.
 XX
 KM Viral replication inhibitor; IRES initiated translation; LAP;
 KM internal ribosome entry site initiated translation; La antigen protein;
 KM viral life cycle; antiviral therapy; pharmaceutical; veterinary;
 KM agricultural; horticultural; virucide; mouse.
 XX
 OS Mus sp.
 XX
 PN WO200283858-A2.
 XX
 PD 24-OCT-2002.
 XX
 PF 12-APR-2002; 2002MO-US011589.
 XX
 PR 16-APR-2001; 2001US-00836073.
 XX
 PA (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.
 XX
 PI Dasgupta A, Das S, Baidya N;
 DR WPI; 2003-058634/05.
 XX
 PT New compound containing acidic and aromatic amino acids, useful as
 PT antiviral therapy in pharmaceutical, veterinary or
 PT agricultural/horticultural applications.
 XX
 PS Claim 1; Page 15; 19pp; English.
 XX
 CC The present invention relates to peptides and methods of inhibiting the
 CC replication of viruses that utilize internal ribosome entry site (IRES)
 CC initiated translation, and/or inhibiting viruses that utilize the La
 CC antigen protein (LAP) in any phase of their life cycle. The peptides of
 CC the invention compete with LAP and inhibit the utilisation of various
 CC biochemical and physiological functions of LAP required for a productive
 CC life cycle. The methods and compositions are useful as antiviral therapy
 CC in pharmaceutical, veterinary or agricultural/horticultural applications.
 CC ABG72101-ABG72119 represent peptides useful as antiviral agents
 CC
 XX
 SQ Sequence 17 AA;
 XX

Query Match 76.0%; Score 76; DB 6; Length 17;
 Best Local Similarity 81.2%; Pred. No. 6.5e-05;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 LDTPKICEQIEYFGDF 18
 |||||
 DB 2 LEAKICHQIEYFGDF 17

RESULT 3
 AAY52200
 ID AAY52200 standard; peptide; 18 AA.
 XX
 AC AAY52200;

XX
 DT 14-MAR-2000 (first entry)
 XX
 DE Human La autoantigen peptide (LAP).
 XX
 KM La autoantigen; LAP; internal ribosome entry site; IRES; translation;
 KM viral replication; IRNA; antiviral agent; picornavirus; flavivirus;
 KM coronavirus; hepatitis virus; rhadovirus; adenovirus; coxsackie virus;
 KM parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
 KM foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
 KM vesicular stomatitis virus.
 XX
 OS Homo sapiens.
 XX
 PN WO9961613-A2.
 XX
 PD 02-DEC-1999.
 XX
 PF 21-MAY-1999; 99WO-US011281.
 XX
 PR 22-MAY-1998; 98US-0086527P.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Das S, Dasgupta A;
 DR WPI; 2000-062712/05.
 XX
 PT New yeast inhibitory peptide useful for inhibiting viral protein
 PT translation and replication.
 XX
 PS Claim 5; Page 57; 81pp; English.
 XX
 CC This sequence is the La autoantigen binding domain (LAP). LAP is a
 CC cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 CC translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors from
 CC binding at the site. The peptide can be used alone or in combination with
 CC an inhibitor RNA (IRNA see AA245200). The LAP peptide is useful as an
 CC antiviral agent, which works through the inhibition of mRNA translation,
 CC especially viral RNA. Examples of viruses which can be inhibited are
 CC picornavirus, flavivirus, coronavirus, hepatitis A B or C viruses,
 CC rhadovirus, adenovirus, and parainfluenza virus, poliovirus, rhinovirus,
 CC coxsackie virus, encephalomyocarditis virus, foot-and-mouth disease
 CC virus, echo virus, hepatitis C virus, infectious bronchitis virus, duck
 CC and human hepatitis B virus, and vesicular stomatitis virus. The peptide
 CC also inhibits replication of the above viruses. The LAP peptide
 CC selectively inhibits viral protein translation, and is therefore not
 CC toxic to the host cell
 CC
 XX
 SQ Sequence 18 AA;
 XX

Query Match 76.0%; Score 76; DB 3; Length 18;
 Best Local Similarity 81.2%; Pred. No. 7e-05;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 LDTPKICEQIEYFGDF 18
 |||||
 DB 3 LEAKICHQIEYFGDF 18

RESULT 4
 ABG72101
 ID ABG72101 standard; peptide; 18 AA.
 XX
 AC ABG72101;
 XX
 DT 28-JAN-2003 (first entry)
 XX
 DE Viral replication inhibiting peptide, LAP.
 XX
 KM Viral replication inhibitor; IRES initiated translation; LAP;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 16:58:20 / Search time 56.3129 Seconds
(without alignments)
90.314 Million cell updates/sec

Title: US-09-836-073-15

Perfect score: 100

Sequence: 1 LDLPKICEQIEYFGDF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29yand4:*

1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	18	6	ABG72115
2	76	76.0	17	6	ABG72113
3	76	76.0	18	3	AAV52200
4	76	76.0	18	6	ABG72101
5	76	76.0	18	6	ABG72114
6	76	76.0	92	3	AA601351
7	76	76.0	408	5	AA603716
8	76	76.0	408	5	ABE65252
9	76	76.0	408	7	ADDE63991
10	76	76.0	408	7	ADDA46272
11	76	76.0	408	7	ADDE63995
12	76	76.0	408	7	ADDE62859
13	76	76.0	415	7	ADDE63989
14	76	76.0	415	7	ADDE63993
15	76	76.0	439	4	ABG08417
16	76	76.0	460	3	AA858987
17	76	76.0	460	5	ABP41511
18	76	76.0	460	5	ABG72102
19	76	76.0	460	5	ABG72104
20	76	76.0	460	5	ABG72109
21	76	76.0	460	5	ABG72111
22	76	76.0	460	5	ABG72112
23	76	76.0	460	5	ABG72110
24	76	76.0	460	5	ABG72116
25	76	76.0	460	5	ABG72103

26	65	65.0	21	2	AA843394	La/SSB ep
27	64	64.0	18	6	ABG72105	Abg72105 Viral rep
28	61	61.0	18	6	ABG72107	Abg72107 Viral rep
29	60	60.0	18	6	ABG72108	Abg72108 Viral rep
30	55	55.0	18	6	ABG72117	Abg72117 Viral rep
31	54	54.0	16	6	ABG72119	Abg72119 Viral rep
32	54	54.0	381	6	ABU19009	Abu19009 Pathogen
33	54	54.0	391	6	ABM72091	Abm72091 Scaphyloc
34	53	53.0	477	6	ABP70526	Abp70526 Histone d
35	51	51.0	377	6	ABU43510	Abu43510 Protein e
36	48	48.0	390	4	ABR65316	ABR65316 Drosophila
37	47	47.0	913	3	AA847714	AA847714 Arabidops
38	47	47.0	923	3	AA847713	AA847713 Arabidops
39	47	47.0	993	3	AA847712	AA847712 Arabidops
40	46	46.5	412	5	AAU96342	AAU96342 Der HMW-m
41	46	46.5	462	5	AAU96341	AAU96341 Der HMW-m
42	46	46.0	491	5	ABR43683	ABR43683 Flj11196
43	46	46.0	492	3	ABR43110	ABR43110 Human ORF
44	46	46.0	706	6	ABU18600	Abu18600 Protein e
45	45	45.0	18	6	ABG72106	Abg72106 Viral rep

ALIGNMENTS

RESULT 1	ABG72115	standard; peptide; 18 AA.
ID	ABG72115	
AC	ABG72115	
AC	ABG72115	
DT	28-JAN-2003	(first entry)
XX		
DE	Viral replication inhibiting peptide, XENOPUS.	
XX		
KW	Viral replication inhibitor; IRBS initiated translation; LAP;	
KW	internal ribosome entry site initiated translation; La antigen protein;	
KW	viral life cycle; antiviral therapy; pharmaceutical; veterinary;	
KW	agricultural; horticultural; virucide; xenopus.	
XX		
OS	Xenopus sp.	
XX		
PN	WO200283858-A2.	
XX		
PD	24-OCT-2002.	
XX		
PF	12-APR-2002; 2002WO-US011589.	
XX		
PR	16-APR-2001; 2001US-00836073.	
XX		
XX	(UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.	
PA	Dasgupta A, Das S, Baidya N;	
XX		
PI	Dasgupta A, Das S, Baidya N;	
XX		
DR	WPI; 2003-058634/05.	
XX		
PT	New compound containing acidic and aromatic amino acids, useful as	
PT	antiviral therapy in pharmaceutical, veterinary or	
PT	agricultural/horticultural applications.	
XX		
PS	Claim 1; Page 15; 19pp; English.	
XX		
CC	The present invention relates to peptides and methods of inhibiting the	
CC	replication of viruses that utilize internal ribosome entry site (IRES)	
CC	initiated translation, and/or inhibiting viruses that utilize the La	
CC	antigen protein (LAP) in any phase of their life cycle. The peptides of	
CC	the invention compete with LAP and inhibit the utilisation of various	
CC	biochemical and physiological functions of LAP required for a productive	
CC	life cycle. The methods and compositions are useful as antiviral therapy	
CC	in pharmaceutical, veterinary or agricultural/horticultural applications.	
CC	ABG72101-ABG72119 represent peptides useful as antiviral agents	
XX		
SO	Sequence 18 AA;	

This Page Blank (uspto)

Mon Sep 13 09:36:11 2004

us-09-836-073-14.rapb

Page 5

Search completed: September 10, 2004, 18:11:55
Job time : 40.424 secs

```
FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-10

Query Match          90.9%; Score 90; DB 9; Length 18;
Best Local Similarity 94.4%; Pred. No. 2.8e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYYFGDF 18
   |||||
Db 1 AALEAKICHQIEYYGDF 18

RESULT 12
US-09-836-073-16
; Sequence 16, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Rat
US-09-836-073-16

Query Match          89.4%; Score 88.5; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 5.1e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 18
   |||||
Db 1 AALEAKICHQIEYYGDF 19

RESULT 13
US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
```

```
US-09-836-073-2

Query Match          87.9%; Score 87; DB 9; Length 18;
Best Local Similarity 88.9%; Pred. No. 8.2e-07;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYYFGDF 18
   |||||
Db 1 AALEAKICHQIEYYGDF 18

RESULT 14
US-09-836-073-3
; Sequence 3, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-3

Query Match          87.9%; Score 87; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 8.2e-07;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYYFGDF 18
   |||||
Db 1 AALEAKICHQIEYYGDF 18

RESULT 15
US-09-836-073-4
; Sequence 4, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-4

Query Match          87.9%; Score 87; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHQIEYYFGDF 18
   |||||
Db 4 EAKICHQIEYYGDF 18
```

TYPE: PRT
ORGANISM: Homo sapiens
US-10-264-049-2643

Query Match 100.0%; Score 99; DB 15; Length 460;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYFGDF 18
|||||
DB 63 AALEAKICHQIEYFGDF 80

RESULT 7
US-09-836-073-13
Sequence 13, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Das, S.
APPLICANT: Balidya, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 220002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13
LENGTH: 17
TYPE: PRT
ORGANISM: Mouse
US-09-836-073-13

Query Match 96.0%; Score 95; DB 9; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.3e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AALEAKICHQIEYFGDF 18
|||||
DB 1 AALEAKICHQIEYFGDF 17

RESULT 8
US-09-836-073-9
Sequence 9, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Das, S.
APPLICANT: Balidya, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 220002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 18
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-836-073-9

Query Match 93.9%; Score 93; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.3e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYFGDF 17
|||||
DB 1 AALEAKICHQIEYFGDF 17

RESULT 9
US-09-836-073-11
Sequence 11, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Das, S.
APPLICANT: Balidya, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 220002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 18
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-836-073-11

Query Match 91.9%; Score 91; DB 9; Length 18;
Best Local Similarity 94.4%; Pred. No. 1.9e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYFGDF 18
|||||
DB 1 AALEAKICHQIEYFGDF 18

RESULT 10
US-09-836-073-12
Sequence 12, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Das, S.
APPLICANT: Balidya, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 220002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 18
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match 91.9%; Score 91; DB 9; Length 18;
Best Local Similarity 94.4%; Pred. No. 1.9e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYFGDF 18
|||||
DB 1 AALEAKICHQIEYFGDF 18

RESULT 11
US-09-836-073-10
Sequence 10, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Das, S.
APPLICANT: Balidya, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

```
APPLICANT: Baidya, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 220002034822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 18
TYPE: PRT
ORGANISM: Bovine
US-09-836-073-14

Query Match      100.0%; Score 99; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1,1e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Oy 1 AALEAKICHQIEYYFGDF 18
Db 1 AALEAKICHQIEYYFGDF 18

RESULT 3
US-10-170-385-477
Sequence 477, Application US/10170385
Publication No. US20030203372A1
GENERAL INFORMATION:
APPLICANT: Ward, Neil Raymond
APPLICANT: Mundy, Christopher Robert
APPLICANT: Kan, On
APPLICANT: Harris, Robert Alan
APPLICANT: White, Jonathan
APPLICANT: Bingley, Katie Mary
APPLICANT: Rayner, William Nigel
APPLICANT: Naylor, Stuart
APPLICANT: Kingman, Susan Mary
APPLICANT: Kirge, David
TITLE OF INVENTION: ANALYSIS METHOD
FILE REFERENCE: 532682000100
CURRENT APPLICATION NUMBER: US/10/170,385
CURRENT FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: PCT/GB02/01662
PRIOR FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: PCT/GB01/05458
PRIOR FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 549
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 477
LENGTH: 408
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-170-385-477

Query Match 100.0%; Score 99; DB 12; Length 408;
Best Local Similarity 100.0%; Pred. No. 2,5e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AALEAKICHQIEYYFGDF 18
Db 11 AALEAKICHQIEYYFGDF 28

RESULT 4
US-09-925-298-695
Sequence 695, Application US/09925298
Publication No. US20020039764A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA103
CURRENT APPLICATION NUMBER: US/09/925,298

```
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05881
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 846
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 695
LENGTH: 460
TYPE: PRT
ORGANISM: Homo sapiens
US-09-925-298-695

Query Match      100.0%; Score 99; DB 12; Length 460;
Best Local Similarity 100.0%; Pred. No. 2,8e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Oy 1 AALEAKICHQIEYYFGDF 18
Db 63 AALEAKICHQIEYYFGDF 80

RESULT 5
US-10-102-806-695
Sequence 695, Application US/10102806
Publication No. US20030054421A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA103P1C1
CURRENT APPLICATION NUMBER: US/10/102,806
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/925,298
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05881
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 846
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 695
LENGTH: 460
TYPE: PRT
ORGANISM: Homo sapiens
US-10-102-806-695

Query Match 100.0%; Score 99; DB 14; Length 460;
Best Local Similarity 100.0%; Pred. No. 2,8e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AALEAKICHQIEYYFGDF 18
Db 63 AALEAKICHQIEYYFGDF 80

RESULT 6
US-10-264-049-2643
Sequence 2643, Application US/10264049
Publication No. US20040005579A1
GENERAL INFORMATION:
APPLICANT: Birse et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA133P1
CURRENT APPLICATION NUMBER: US/10/264,049
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/18569
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: US 60/209,467
PRIOR FILING DATE: 2000-06-07
NUMBER OF SEQ ID NOS: 4360
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 2643
LENGTH: 460

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:52:06; Search time 40.324 Seconds

(without alignments)
143.151 Million cell updates/sec

Title: US-09-836-073-14

Perfect score: 99

Sequence: 1 AALEAKICHOIEYFGDF 18

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	100.0	18	9	US-09-836-073-1
2	99	100.0	18	9	US-09-836-073-14
3	99	100.0	408	12	US-10-170-385-477
4	99	100.0	460	12	US-09-925-298-695
5	99	100.0	460	14	US-10-102-806-695
6	99	100.0	460	15	US-10-264-049-2643
7	95	96.0	17	9	US-09-836-073-13
8	93	93.9	18	9	US-09-836-073-9
9	91	91.9	18	9	US-09-836-073-11
10	91	91.9	18	9	US-09-836-073-12
11	90	90.9	18	9	US-09-836-073-10
12	88.5	89.4	19	9	US-09-836-073-16
13	87	87.9	18	9	US-09-836-073-2
14	87	87.9	18	9	US-09-836-073-3
15	87	87.9	18	9	US-09-836-073-4

16	84.8	84.8	18	9	US-09-836-073-7	Sequence 7, Appli
17	83	83.8	18	9	US-09-836-073-8	Sequence 8, Appli
18	76	76.8	18	9	US-09-836-073-15	Sequence 15, Appli
19	75	75.8	18	9	US-09-836-073-5	Sequence 5, Appli
20	73	73.7	21	15	US-10-376-121A-20	Sequence 20, Appli
21	68	68.7	18	9	US-09-836-073-6	Sequence 6, Appli
22	68	68.7	38	12	US-10-325-810-214	Sequence 21, App
23	59	59.6	38	9	US-09-843-676-25	Sequence 25, Appl
24	59	59.6	38	9	US-09-766-253-25	Sequence 25, Appl
25	59	59.6	38	10	US-09-438-486-25	Sequence 25, Appl
26	59	59.6	38	12	US-10-325-810-215	Sequence 25, App
27	59	59.6	38	14	US-10-053-758-25	Sequence 25, Appl
28	59	59.6	38	14	US-10-054-295-25	Sequence 25, Appl
29	59	59.6	38	14	US-10-054-611-25	Sequence 25, Appl
30	58.5	59.1	37	9	US-09-843-676-24	Sequence 24, Appl
31	58.5	59.1	37	9	US-09-766-253-24	Sequence 24, Appl
32	58.5	59.1	37	10	US-09-438-486-24	Sequence 24, Appl
33	58.5	59.1	37	14	US-10-053-758-24	Sequence 24, Appl
34	58.5	59.1	37	14	US-10-054-295-24	Sequence 24, Appl
35	58.5	59.1	37	14	US-10-054-611-24	Sequence 24, Appl
36	57	57.6	16	9	US-09-836-073-19	Sequence 19, Appl
37	53	53.5	922	16	US-10-437-963-125036	Sequence 125036,
38	52	52.5	420	16	US-10-437-963-128072	Sequence 128072,
39	50	50.5	143	12	US-10-424-559-254661	Sequence 254661,
40	50	50.5	395	12	US-10-424-559-254664	Sequence 254664,
41	50	50.5	411	14	US-10-177-478-8	Sequence 8, Appli
42	48	48.5	39	9	US-09-843-676-26	Sequence 26, Appl
43	48	48.5	39	9	US-09-766-253-26	Sequence 26, Appl
44	48	48.5	39	10	US-09-438-486-26	Sequence 26, Appl
45	48.5	48.5	39	12	US-10-325-810-216	Sequence 216, App

ALIGNMENTS

```
US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-1

Query Match      100.0%; Score 99; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AALEAKICHOIEYFGDF 18
Db      1 AALEAKICHOIEYFGDF 18

RESULT 2
US-09-836-073-14
; Sequence 14, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-24

Query Match 59.1%; Score 58.5; DB 3; Length 37;
Best Local Similarity 91.7%; Pred. No. 0.0075;
Matches 11; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 7 ICHQIEYFGDF 18
Db 1 ICHQ-EYFGDF 11

RESULT 15
US-09-430-323-24
Sequence 24, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Langner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin
Andrews, William H.
TITLE OF INVENTION: NO. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-430-323-24

Query Match 59.1%; Score 58.5; DB 4; Length 37;
Best Local Similarity 91.7%; Pred. No. 0.0075;
Matches 11; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 7 ICHQIEYFGDF 18
Db 1 ICHQ-EYFGDF 11

Search completed: September 10, 2004, 18:05:09
Job time: 15.8883 secs

FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 215:
US-09-721-456-215

Query Match 59.6%; Score 59; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0064; 1;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIEYFGD 17
DB 1 ICHQIEYFGD 11

RESULT 13
US-08-851-843A-24
Sequence 24, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-24

Query Match 59.1%; Score 58.5; DB 3; Length 37;
Best Local Similarity 91.7%; Pred. No. 0.0075; 0;
Matches 11; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 7 ICHQIEYFGDF 18
DB 1 ICHQIEYFGDF 11

RESULT 14
US-08-854-050-24
Sequence 24, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6261836el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536

FILED DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-430-323-25

Query Match 59.6%; Score 59; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0064;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ICHOIEYFFGD 17
Db 1 ICEOIEYFFGD 11

RESULT 11
US-09-402-181B-215
Sequence 215, Application US/09402181B
Patent No. 6610839
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 633
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,181B
FILING DATE: 29-Sep-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050

FILED DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ausenhue, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 215:
US-09-402-181B-215

Query Match 59.6%; Score 59; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0064;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ICHOIEYFFGD 17
Db 1 ICEOIEYFFGD 11

RESULT 12
US-09-721-456-215
Sequence 215, Application US/09721456
Patent No. 6617110
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/721,456
FILING DATE: 22-NOV-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017

FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-215

Query Match 59.6%; Score 59; DB 3; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0064;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIEYFGD 17
DB 1 ICEQIEYFGD 11

RESULT 9
US-08-854-050-25
Sequence 25, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6261836el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-25

Query Match 59.6%; Score 59; DB 3; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0064;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIEYFGD 17
DB 1 ICEQIEYFGD 11

RESULT 10
US-09-430-323-25
Sequence 25, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843

FILED DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: Peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 214:
US-09-721-456-214

Query Match 68.7%; Score 68; DB 4; Length 38;
Best Local Similarity 91.7%; Pred. No. 0.00022;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 7 ICHQIEYFGDF 18
Db 1 ICHQIEYFGDF 12

RESULT 7
US-08-851-843A-25
Sequence 25, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cecch, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Hatley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809e1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-25

Query Match 59.6%; Score 59; DB 3; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0064;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 7 ICHQIEYFGDF 17
Db 1 ICHQIEYFGDF 11

RESULT 8
US-08-974-549A-215
Sequence 215, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cecch, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Hatley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050

APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0300
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-214

Query Match 68.7%; Score 68; DB 3; Length 38;
Best Local Similarity 91.7%; Pred. No. 0.00022;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 7 ICHQIEYYFGDF 18
Db 1 ICHQIEYYFGDF 12

RESULT 5
US-09-402-181B-214
Sequence 214, Application US/09402181B
Patent No. 6610839
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 633
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,181B
FILING DATE: 29-Sep-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503

FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Auehbus, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0300
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 214:
US-09-402-181B-214

Query Match 68.7%; Score 68; DB 4; Length 38;
Best Local Similarity 91.7%; Pred. No. 0.00022;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 7 ICHQIEYYFGDF 18
Db 1 ICHQIEYYFGDF 12

RESULT 6
US-09-721-456-214
Sequence 214, Application US/09721456
Patent No. 6617110
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/721,456
FILING DATE: 22-NOV-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312

PRIOR FILING DATE: 1997-10-06
PRIOR APPLICATION NUMBER: 08/321,427
PRIOR FILING DATE: 1994-10-11
PRIOR APPLICATION NUMBER: 60/086,527
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 4
LENGTH: 18
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: B-LAP
US-09-316-630-4

Query Match 100.0%; Score 99; DB 3; Length 18;
Best Local Similarity 100.0%; Pred. No. 8,5e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AALEAKICHQIEYFGDF 18
DB 1 AALEAKICHQIEYFGDF 18

RESULT 3
US-08-475-955-20
Sequence 20, Application US/08475955
Patent No. 6641813
GENERAL INFORMATION:
APPLICANT: Harley, John
TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS OF
NUMBER OF SEQUENCES: 218
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Patricia L. Pabst
STREET: 2800 One Atlantic Center, 1250 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,955
FILING DATE: June 7, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/867,819
FILING DATE: April 13, 1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/648,205
FILING DATE: January 31, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/472,947
FILING DATE: January 31, 1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patricia L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRF14CIP(2) DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-873-8795
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Binding-site
LOCATION: 8..15
US-08-475-955-20

Query Match 73.7%; Score 73; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 ICHOIEYFYFGDF 18
DB 1 ICHOIEYFYFGDF 12

RESULT 4
US-08-974-549A-214
Sequence 214, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:43:01; Search time 15.8883 Seconds

(without alignments)
58.488 Million cell updates/sec

Title: US-09-836-073-14

Perfect score: 99

Sequence: 1 AALEAKICHOIEYFGDF 18

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	100.0	18	3	US-09-316-630-3
2	99	100.0	18	3	US-09-316-630-4
3	73	73.7	21	4	US-08-475-955-20
4	68	68.7	38	3	US-08-974-549A-214
5	68	68.7	38	4	US-09-402-181B-214
6	68	68.7	38	4	US-09-721-456-214
7	59	59.6	38	3	US-08-851-843A-25
8	59	59.6	38	3	US-08-974-549A-215
9	59	59.6	38	3	US-08-854-050-25
10	59	59.6	38	4	US-09-430-323-25
11	59	59.6	38	4	US-09-402-181B-215
12	59	59.6	38	4	US-09-721-456-215
13	58.5	59.1	37	3	US-08-851-843A-24
14	58.5	59.1	37	3	US-08-854-050-24
15	58.5	59.1	37	4	US-09-430-323-24
16	48	48.5	39	3	US-08-851-843A-26
17	48	48.5	39	3	US-08-974-549A-216
18	48	48.5	39	3	US-08-854-050-26
19	48	48.5	39	3	US-09-430-323-26
20	48	48.5	39	4	US-09-402-181B-216
21	48	48.5	39	4	US-09-721-456-216
22	42	42.4	456	1	US-08-205-719-4
23	42	42.4	456	3	US-08-431-517F-6
24	42	42.4	482	3	US-08-431-517F-5
25	42	42.4	482	6	5245013-2
26	41.5	41.9	124	4	US-09-489-039A-9103
27	41	41.4	483	3	US-09-027-166-7

28	41	41.4	497	4	US-09-489-039A-8926	Sequence 8926, Ap
29	41	41.4	1220	2	US-08-680-326-38	Sequence 38, Ap1
30	40	40.4	431	1	US-08-311-023-2	Sequence 2, Ap1
31	40	40.4	593	4	US-09-252-991A-23340	Sequence 23340, A
32	40	40.4	868	4	US-09-800-729-106	Sequence 106, App
33	40	40.4	921	4	US-09-800-729-199	Sequence 199, App
34	39	39.4	69	4	US-09-621-976-7385	Sequence 7385, Ap
35	39	39.4	582	4	US-09-976-594-733	Sequence 733, App
36	39	39.4	754	2	US-08-941-262-1	Sequence 1, Ap1
37	39	39.4	755	2	US-08-941-262-3	Sequence 3, Ap1
38	39	39.4	811	4	US-09-199-617A-93	Sequence 93, Ap1
39	39	39.4	812	1	US-08-248-628A-1	Sequence 1, Ap1
40	39	39.4	812	1	US-08-451-932-1	Sequence 1, Ap1
41	39	39.4	812	1	US-08-452-260-1	Sequence 1, Ap1
42	39	39.4	812	1	US-08-326-785-1	Sequence 1, Ap1
43	39	39.4	812	2	US-08-612-788-1	Sequence 1, Ap1
44	39	39.4	812	2	US-08-605-598B-1	Sequence 1, Ap1
45	39	39.4	812	2	US-08-429-743-1	Sequence 1, Ap1

ALIGNMENTS

```
RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548-21
; CURRENT APPLICATION NUMBER: US/09/316,630
; PRIOR FILING DATE: 1997-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3
Query Match          100.0%; Score 99; DB 3; Length 18;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB          1 AALEAKICHOIEYFGDF 18
           1 AALEAKICHOIEYFGDF 18
RESULT 2
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548-21
; CURRENT APPLICATION NUMBER: US/09/316,630
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
```

This Page Blank (uspto)

XX 14-AUG-2002; 2002WO-US025765.
 PF 14-AUG-2001; 2001US-0312147P.
 XX 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX (GENO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 PI Woolf C, D'urso D, Befort K, Costigan M,
 XX WPI: 2003-268312/26.
 DR GENBANK; P38656.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX
 PS Claim 1; Page; 1017pp; English.
 XX
 CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 415 AA;
 XX
 Query Match 100.0%; Score 99; DB 7; Length 415;
 Best Local Similarity 100.0%; Pred. No. 2.8e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AALEAKICHQIEYFGDF 18
 ||||||||||||||||
 DB 11 AALEAKICHQIEYFGDF 28
 ||||||||||||||||
 RESULT 13
 AAG08417
 ID AAG08417 standard; protein; 439 AA.
 XX
 AC AAG08417;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #8408.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.

XX WO200175067-A2.
 PN 11-OCT-2001.
 XX
 PD 30-MAR-2001; 2001WO-US008631.
 XX
 PF 31-MAR-2000; 2000US-00540217.
 XX 23-AUG-2000; 2000US-00649167.
 PR (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 PI WPI: 2001-639362/73.
 XX N-PSDB; AAS72604.
 DR
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 38776; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAG0010-AAG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 439 AA;
 XX
 Query Match 100.0%; Score 99; DB 4; Length 439;
 Best Local Similarity 100.0%; Pred. No. 3e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AALEAKICHQIEYFGDF 18
 ||||||||||||||||
 DB 41 AALEAKICHQIEYFGDF 58
 ||||||||||||||||
 RESULT 14
 AAB58987
 ID AAB58987 standard; protein; 460 AA.
 XX
 AC AAB58987;
 XX
 DT 27-MAR-2001 (first entry)
 XX
 DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 695.
 XX
 KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW neutropic; neuroprotective; antiviral; antiallergic; hepatocytic;
 KW antidiabetic; antiinflammatory; anticancer; antiviral; anticonvulsant;
 KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;

PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX WPI, 2003-268312/26.
DR GENBANK, P05455.
XX
PT New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page: 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC mediates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
CC
XX
SQ Sequence 408 AA;
CC
Query Match 100.0%; Score 99; DB 7; Length 408;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AALEAKICHQIEYYFGDF 18
DB |||||
11 AALEAKICHQIEYYFGDF 28
XX
RESULT 11
ADE63989
ID ADE63989 standard; protein; 415 AA.
XX
AC ADE63989;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein P38656, SEQ ID NO 9935.
XX
KM Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
OS Rattus norvegicus.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.
PR, 26-NOV-2001; 2001US-0333347P.
XX
XX (GENO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
PI WPI, 2003-268312/26.
DR GENBANK, P38656.
XX
PT New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page: 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC mediates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
CC
XX
SQ Sequence 415 AA;
CC
Query Match 100.0%; Score 99; DB 7; Length 415;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AALEAKICHQIEYYFGDF 18
DB |||||
11 AALEAKICHQIEYYFGDF 28
XX
RESULT 12
ADE63993
ID ADE63993 standard; protein; 415 AA.
XX
AC ADE63993;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein P38656, SEQ ID NO 9939.
XX
KM Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
OS Rattus norvegicus.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.

PS Claim 1, Page, 1017pp, English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 408 AA;

Query Match 100.0%; Score 99; DB 7; Length 408;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALEAKICHOIEYFGDF 18
| | | | | | | | | | | | | | | | | | | | | |
DB 11 AALEAKICHOIEYFGDF 28

RESULT 9
ADE63995
ID ADE63995 standard; protein; 408 AA.
XX
AC ADE63995;
XX

DT 29-JAN-2004 (first entry)
XX
DE Human Protein P05455, SEQ ID NO 9941.

XX Human; pain; neuronal tissue; gene therapy;
XX spinal segmental nerve injury; chronic constriction injury; CCI;
XX spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.

XX
XX WO2003016475-A2.
XX
PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.

XX
XX (GEHO) GEN HOSPITAL CORP.
XX (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR

DR GENBANK; P05455.

XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
PT

PS Claim 1, Page, 1017pp, English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 408 AA;

Query Match 100.0%; Score 99; DB 7; Length 408;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALEAKICHOIEYFGDF 18
| | | | | | | | | | | | | | | | | | | | | |
DB 11 AALEAKICHOIEYFGDF 28

RESULT 10
ADE62859
ID ADE62859 standard; protein; 408 AA.
XX
AC ADE62859;
XX

DT 29-JAN-2004 (first entry)
XX
DE Human Protein P05455, SEQ ID NO 8793.

XX Human; pain; neuronal tissue; gene therapy;
XX spinal segmental nerve injury; chronic constriction injury; CCI;
XX spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.

XX
XX WO2003016475-A2.
XX
PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.

XX
XX (GEHO) GEN HOSPITAL CORP.
XX (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR

PS Claim 35; Page 425; 538pp; English.

XX The present invention relates to methods for identifying genes and
CC proteins that are implicated in a specific disease or physiological
CC condition. The method comprises comparing the transcriptome/proteome of a
CC specialised cell type implicated in a disease or condition with that of a
CC second specialised cell type, under two experimental conditions, and
CC identifying a gene that is differentially regulated in the two
CC specialised cell types under experimental conditions. ABV7873-ABV78116
CC and ABP65061-ABP65257 were identified using the methods of the invention.
CC The coding sequences and proteins are useful for treating a disease in a
CC patient, for manufacture of a medicament for treating hypoxia-regulated
CC conditions, and for regulating tumorigenesis, angiogenesis, apoptosis,
CC biological response to hypoxia conditions, or hypoxic-associated
CC pathology in a patient. The coding sequences and proteins are also useful
CC for monitoring the therapeutic treatment of a disease or physiological
CC condition, such as cancer, ischemic conditions, reperfusion injury,
CC retinopathy, neonatal stress, preclampsia, atherosclerosis, inflammatory
CC conditions, wound healing, inflammation, erythropoiesis or hair loss

XX Sequence, 408 AA;

SO

Query Match 100.0%; Score 99; DB 5; Length 408;
Best Local Similarity 100.0%; Pred. No. 2,8e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AALEAKICHQIEYFGDF 18
|||
DB 11 AALEAKICHQIEYFGDF 28

RESULT 7
ADE63991
ID ADE63991 standard; protein; 408 AA.

XX ADE63991;
AC

DT 29-JAN-2004 (first entry)

XX Human Protein P05455, SEQ ID NO 9937.

DE

XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KM spared nerve injury; SNI; Chung.

XX Homo sapiens.

OS

PN WO2003016475-A2.

PD 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

PF

XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.

XX (GHEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;
PI WPI; 2003-268312/26.
DR GENBANK; P05455.

XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also

CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC- which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 408 AA;

SO

Query Match 100.0%; Score 99; DB 7; Length 408;
Best Local Similarity 100.0%; Pred. No. 2,8e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AALEAKICHQIEYFGDF 18
|||
DB 11 AALEAKICHQIEYFGDF 28

RESULT 8
ADD46272
ID ADD46272 standard; protein; 408 AA.

XX ADD46272;
AC

DT 29-JAN-2004 (first entry)

XX Human Protein P05455, SEQ ID NO 11947.

DE

XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KM spared nerve injury; SNI; Chung.

XX Homo sapiens.

OS

PN WO2003016475-A2.

PD 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

PF

XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.

XX (GHEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;
PI WPI; 2003-268312/26.
DR GENBANK; P05455.

XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.

XX

KW gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-00200610.
 XX
 PR 26-FEB-1999; 99US-0122487P.
 XX
 PA (GENEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-500381/45.
 XX
 DR N-PSDB; AAC01357.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
 XX
 PS Claim 13; SEQ ID NO 5432; 71bp + Sequence Listing; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number of
 CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
 CC prepared from total human RNAs or polyA+ RNAs derived from 30 different
 CC tissues. EST sequences usually correspond mainly to the 3' untranslated
 CC region (UTR) of the mRNA because they are often obtained from oligo-dT
 CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
 CC sequences derived from the 5' ends of mRNAs and even in those cases where
 CC longer cDNA sequences have been obtained, the full 5' UTR is rarely
 CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can
 CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
 CC are also used in diagnostic, forensic, gene therapy and chromosome
 CC mapping procedures. They are used to obtain upstream regulatory sequences
 CC and to design expression and secretion vectors
 XX
 SQ Sequence 92 AA;
 XX
 Query Match 100.0%; Score 99; DB 3; Length 92;
 Best Local Similarity 100.0%; Pred. No. 5.8e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 AALEAKICHQIEYFGDF 18
 |||||
 DB 11 AALEAKICHQIEYFGDF 28
 XX
 RESULT 5
 AAM03716
 ID AAM03716 standard; protein; 408 AA.
 XX
 AC AAM03716;
 XX
 DT 25-MAR-2003 (revised)
 DT 12-MAR-1997 (first entry)
 XX
 DE Human autoantigen Ia(SS-B).
 XX
 KW Autoimmune disease; Ia autoantigen; Sjogren's syndrome;
 KW systemic lupus erythematosus; diagnosis.
 XX
 OS Homo sapiens.
 XX
 PN US5541291-A.
 XX
 PD 30-JUL-1996.
 XX
 PF 27-MAY-1987; 87US-00054871.
 XX
 PR 31-DEC-1984; 84US-00687908.
 XX

XX
 PA (UYDU-) UNIV DUKE.
 XX
 PI Keene JD;
 XX
 DR WPI; 1996-362015/36.
 XX
 PT Auto-antigen U2-RNP, associated with severe polymyositis-scleroderma
 PT overlap syndrome - useful for diagnosis and treatment of autoimmune
 PT diseases.
 XX
 PS Disclosure; Col 15-16; 21pp; English.
 XX
 CC The human lupus antigen (Ia) is diagnostic for Sjogren's syndrome, as
 CC well as occurring in systemic lupus erythematosus patients. The Ia protein
 CC is clinically related to the Ro protein that is highly common among
 CC autoimmune patients. Ia and Ro antigens sometimes reside on the same
 CC cellular ribonucleoprotein particle; most Ia patients contain some Ro
 CC antibodies and vice versa. Ia cDNA has been isolated from a human liver
 CC library. (Updated on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 408 AA;
 XX
 Query Match 100.0%; Score 99; DB 2; Length 408;
 Best Local Similarity 100.0%; Pred. No. 2.8e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 AALEAKICHQIEYFGDF 18
 |||||
 DB 11 AALEAKICHQIEYFGDF 28
 XX
 RESULT 6
 ABP65252
 ID ABP65252 standard; protein; 408 AA.
 XX
 AC ABP65252;
 XX
 DT 12-NOV-2002 (first entry)
 XX
 DE Hypoxia-regulated protein #126.
 XX
 KW Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;
 KW antiinflammatory; vulnerrary; gynecological; ophthalmological; vaccine;
 KW hypoxia; tumourigenesis; angiogenesis; apoptosis; cancer;
 KW ischaemic condition; reperfusion injury; retinopathy; neonatal stress;
 KW preclapmsia; atherosclerosis; inflammatory condition; wound healing;
 KW inflammation; erythropoiesis; hair loss; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200246465-A2.
 XX
 PD 13-JUN-2002.
 XX
 PF 10-DEC-2001; 2001WO-GB005458.
 XX
 PR 08-DEC-2000; 2000GB-00030076.
 PR 08-FEB-2001; 2001GB-00003156.
 PR 25-OCT-2001; 2001GB-00025666.
 XX
 PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 PI White J, Mundy CR, Ward NR, Krige D, Kingsman SM, Harris RA;
 PI Rayner WN;
 XX
 DR WPI; 2002-627238/67.
 XX
 PT Identifying a gene involved in disease for treating hypoxia-regulated
 PT conditions, comprises comparing the transcriptome/proteome of two cell
 PT types under different conditions and identifying a differentially
 PT regulated gene.
 XX

CC rhabdovirus, adenovirus, and parainfluenza virus, poliovirus, rhinovirus,
 CC coxsackie virus, encephalomyocarditis virus, foot-and-mouth disease
 CC virus, echo virus, hepatitis C virus, infectious bronchitis virus, duck
 CC and human hepatitis B virus, and vesicular stomatitis virus. The peptide
 CC also inhibits replication of the above viruses. The LAP peptide
 CC selectively inhibits viral protein translation, and is therefore not
 CC toxic to the host cell

XX Sequence 18 AA;

Query Match 100.0%; Score 99; DB 3; Length 18;
 Best Local Similarity 100.0%; Pred. No. 1e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AALEAKICHQIEYFGDF 18
 Db 1 AALEAKICHQIEYFGDF 18

RESULT 2

ABG72101
 ID ABG72101 standard; peptide; 18 AA.

XX AC ABG72101;

XX DT 28-JAN-2003 (first entry)

XX DE Viral replication inhibiting peptide, LAP.

XX KM Viral replication inhibitor; IRES initiated translation; LAP;

XX KM internal ribosome entry site initiated translation; La antigen protein;

XX KM viral life cycle; antiviral therapy; pharmaceutical; veterinary;

XX KM agricultural; horticultural; viroicide.

XX OS Unidentified.

XX PN WO200283858-A2.

XX PD 24-OCT-2002.

XX PF 12-APR-2002; 2002WO-US011589.

XX PR 16-APR-2001; 2001US-00836073.

XX PA (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.

XX PI Dasgupta A, Das S, Baidya N;

XX DR WPI; 2003-058634/05.

XX PT New compound containing acidic and aromatic amino acids, useful as
 PT antiviral therapy in pharmaceutical, veterinary or
 PT agricultural/horticultural applications.

XX PS Claim 1; Page 15; 19pp; English.

XX CC The present invention relates to peptides and methods of inhibiting the
 CC replication of viruses that utilise internal ribosome entry site (IRES)
 CC initiated translation, and/or inhibiting viruses that utilise the La
 CC antigen protein (LAP) in any phase of their life cycle. The peptides of
 CC the invention compete with LAP and inhibit the utilisation of various
 CC biochemical and physiological functions of LAP required for a productive
 CC life cycle. The methods and compositions are useful as antiviral therapy
 CC in pharmaceutical, veterinary or agricultural/horticultural applications.
 CC ABG72101-ABG72119 represent peptides useful as antiviral agents

XX SQ Sequence 18 AA;

Query Match 100.0%; Score 99; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 1e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AALEAKICHQIEYFGDF 18

Db 1 AALEAKICHQIEYFGDF 18

RESULT 3

ABG72114
 ID ABG72114 standard; peptide; 18 AA.

XX AC ABG72114;

XX DT 28-JAN-2003 (first entry)

XX DE Viral replication inhibiting peptide, BOVINE.

XX KM Viral replication inhibitor; IRES initiated translation; LAP;

XX KM internal ribosome entry site initiated translation; La antigen protein;

XX KM viral life cycle; antiviral therapy; pharmaceutical; veterinary;

XX KM agricultural; horticultural; viroicide; bovine.

XX OS Bovinae.

XX PN WO200283858-A2.

XX PD 24-OCT-2002.

XX PF 12-APR-2002; 2002WO-US011589.

XX PR 16-APR-2001; 2001US-00836073.

XX PA (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.

XX PI Dasgupta A, Das S, Baidya N;

XX DR WPI; 2003-058634/05.

XX PT New compound containing acidic and aromatic amino acids, useful as
 PT antiviral therapy in pharmaceutical, veterinary or
 PT agricultural/horticultural applications.

XX PS Disclosure; Page 6; 19pp; English.

XX CC The present invention relates to peptides and methods of inhibiting the
 CC replication of viruses that utilise internal ribosome entry site (IRES)
 CC initiated translation, and/or inhibiting viruses that utilise the La
 CC antigen protein (LAP) in any phase of their life cycle. The peptides of
 CC the invention compete with LAP and inhibit the utilisation of various
 CC biochemical and physiological functions of LAP required for a productive
 CC life cycle. The methods and compositions are useful as antiviral therapy
 CC in pharmaceutical, veterinary or agricultural/horticultural applications.
 CC ABG72101-ABG72119 represent peptides useful as antiviral agents

XX SQ Sequence 18 AA;

Query Match 100.0%; Score 99; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 1e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AALEAKICHQIEYFGDF 18
 Db 1 AALEAKICHQIEYFGDF 18

RESULT 4

AAG01351
 ID AAG01351 standard; protein; 92 AA.

XX AC AAG01351;

XX DT 06-OCT-2000 (first entry)

XX DE Human secreted protein, SEQ ID NO: 5432.

XX KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 16:58:20 ; Search time 56.3129 Seconds
(without alignments)
90.314 Million cell updates/sec

Title: US-09-836-073-14

Perfect score: 99
Sequence: 1 AALFAKICHQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp1980s:*\n2: geneseqp1990s:*\n3: geneseqp2000s:*\n4: geneseqp2001s:*\n5: geneseqp2002s:*\n6: geneseqp2003as:*\n7: geneseqp2003bs:*\n8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	100.0	18	3 AAY52200	Aay52200 Human la
2	99	100.0	18	6 ABG72101	Abg72101 Viral rep
3	99	100.0	18	6 ABG72114	Abg72114 Viral rep
4	99	100.0	92	3 AAG01351	Aag01351 Human sec
5	99	100.0	408	2 AAW03716	Aaw03716 Human aut
6	99	100.0	408	5 ABP65252	Abp65252 Hypoxia-r
7	99	100.0	408	7 ADE63991	Ad63991 Human Pro
8	99	100.0	408	7 ADD46272	Add46272 Human Pro
9	99	100.0	408	7 ADE63995	Ad63995 Human Pro
10	99	100.0	408	7 ADE62859	Ad62859 Human Pro
11	99	100.0	415	7 ADE63989	Ad63989 Rat Prote
12	99	100.0	415	7 ADE63993	Ad63993 Rat Prote
13	99	100.0	439	4 ABG08417	Abg08417 Novel hum
14	99	100.0	460	3 AAB58987	Aab58987 Breast an
15	99	100.0	460	3 ABP41511	Abp41511 Human ova
16	95	96.0	17	6 ABG72113	Abg72113 Viral rep
17	93	93.9	18	6 ABG72109	Abg72109 Viral rep
18	91	91.9	18	6 ABG72111	Abg72111 Viral rep
19	91	91.9	18	6 ABG72112	Abg72112 Viral rep
20	90	90.9	18	6 ABG72110	Abg72110 Viral rep
21	88.5	89.4	19	6 ABG72116	Abg72116 Viral rep
22	87	87.9	18	6 ABG72102	Abg72102 Viral rep
23	87	87.9	18	6 ABG72103	Abg72103 Viral rep
24	87	87.9	18	6 ABG72104	Abg72104 Viral rep
25	84	84.8	18	6 ABG72107	Abg72107 Viral rep

26	83	83.8	18	6 ABG72108	Abg72108 Viral rep
27	76	76.8	18	6 ABG72115	Abg72115 Viral rep
28	75	75.8	18	6 ABG72105	Abg72105 Viral rep
29	73	73.7	21	2 AAR43394	Aar43394 La/SSB ep
30	68	68.7	18	6 ABG72106	Abg72106 Viral rep
31	57	57.6	16	6 ABG72119	Abg72119 Viral rep
32	51	51.5	390	4 ABB65316	Abb65316 Drosophila
33	50	50.5	411	6 ABB70526	Abb70526 Histone d
34	49	49.5	381	6 ABB19009	Abb19009 Pathogen
35	49	49.5	391	6 ABB72091	Abb72091 Staphyloc
36	48	48.5	913	3 AAG47714	Aag47714 Arabidops
37	48	48.5	923	3 AAG47712	Aag47712 Arabidops
38	48	48.5	993	3 AAG47711	Aag47711 Arabidops
39	46	46.5	377	6 ABU43510	Abu43510 Protein e
40	44	44.4	18	6 ABG72117	Abg72117 Viral rep
41	44	44.4	135	4 ABG27058	Abg27058 Novel hum
42	44	44.4	224	5 ABB89645	Abb89645 Human pol
43	44	44.4	224	6 ADA55336	Ada55336 Human pro
44	43	43.4	176	4 AAU52458	Aau52458 Propionib
45	43	43.4	176	6 ABM48977	Abm48977 Propionib

ALIGNMENTS

RESULT 1
AAY52200 standard; peptide; 18 AA.

14-MAR-2000 (first entry)

Human la autoantigen peptide (LAP).

La autoantigen; LAP, internal ribosome entry site; IRBS; translation; viral replication; RNA; antiviral agent; picornavirus; flavivirus; coronavirus; hepatitis virus; rhinovirus; adenovirus; coxsackie virus; KM parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus; foot-and-mouth disease virus; echo virus; infectious bronchitis virus; vesicular stomatitis virus.

OS Homo sapiens.

PN WO961613-A2.

PD 02-DEC-1999.

PF 21-MAY-1999; 99WO-US011281.

PR 22-MAY-1998; 98US-0086527P.

PA (REGC) UNIV CALIFORNIA.

PI Das S, Dasgupta A;

DR WPI; 2000-062712/05.

PT New yeast inhibitory peptide useful for inhibiting viral protein translation and replication.

PS Claim 5; Page 57; 81pp; English.

This sequence is the la autoantigen binding domain (LAP). LAP is a cellular protein which binds to an internal ribosome entry site (IRBS). The peptide is used to inhibit mRNA translation, viral protein translation or viral replication. Viral mRNA translation is initiated at the IRBS and the LAP peptide prevents translation initiation factors from binding at the site. The peptide can be used alone or in combination with an inhibitor RNA (RNA see AA245200). The LAP peptide is useful as an antiviral agent, which works through the inhibition of mRNA translation, especially viral mRNA. Examples of viruses which can be inhibited are picornavirus, flavivirus, coronavirus, hepatitis A B or C viruses.

This Page Blank (uspto)

Search completed: September 10, 2004, 18:11:55
Job time : 40.424 secs

```
FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-11

Query Match      85.6%; Score 83; DB 9; Length 18;
Best Local Similarity 88.9%; Pred. No. 1.9e-06;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 AALEAKICHQIEYFGDF 18
   |||||
Db 1 AALEAKICHQIEYFGDF 18

RESULT 12
US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-10

Query Match      84.5%; Score 82; DB 9; Length 18;
Best Local Similarity 88.9%; Pred. No. 2.7e-06;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 AALEAKICHQIEYFGDF 18
   |||||
Db 1 AALEAKICHQIEYFGDF 18

RESULT 13
US-09-836-073-16
; Sequence 16, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Rat
```

```
US-09-836-073-16

Query Match      83.0%; Score 80.5; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 5e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Cy 1 AALEAKICHQIEYFGDF 18
   |||||
Db 1 AALEAKICHQIEYFGDF 19

RESULT 14
US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match      81.4%; Score 79; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 8.2e-06;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Cy 1 AALEAKICHQIEYFGDF 18
   |||||
Db 1 AALEAKICHQIEYFGDF 18

RESULT 15
US-09-836-073-3
; Sequence 3, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-3

Query Match      81.4%; Score 79; DB 9; Length 18;
Best Local Similarity 77.8%; Pred. No. 8.2e-06;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Cy 1 AALEAKICHQIEYFGDF 18
   |||||
Db 1 AALEAKICHQIEYFGDF 18
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-806-695

Query Match
Best Local Similarity 93.8%; Score 91; DB 14; Length 460;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYQFGDF 18
| | | | | | | | | | | | | | | | | |
Db 63 AALEAKICHQIEYQFGDF 80

RESULT 7
US-10-264-049-2643
; Sequence 2643, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA13P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: Patentin Ver. 3.1
; SEQ ID NO 2643
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-2643

Query Match
Best Local Similarity 93.8%; Score 91; DB 15; Length 460;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYQFGDF 18
| | | | | | | | | | | | | | | | | |
Db 63 AALEAKICHQIEYQFGDF 80

RESULT 8
US-09-836-073-8
; Sequence 8, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-073-8

Query Match
Best Local Similarity 91.8%; Score 89; DB 9; Length 18;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYQFGDF 18
| | | | | | | | | | | | | | | | | |
Db 1 AALEAKICHQIEYQFGDF 18
```

```

RESULT 9
US-09-836-073-13
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Mouse
US-09-836-073-13

Query Match
Best Local Similarity 89.7%; Score 87; DB 9; Length 17;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AALEAKICHQIEYQFGDF 18
| | | | | | | | | | | | | | | | | |
Db 1 AALEAKICHQIEYQFGDF 17

RESULT 10
US-09-836-073-9
; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-073-9

Query Match
Best Local Similarity 87.6%; Score 85; DB 9; Length 18;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYQFGDF 17
| | | | | | | | | | | | | | | | | |
Db 1 AALEAKICHQIEYQFGDF 17

RESULT 11
US-09-836-073-11
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
```

```

; APPLICANT: Baldeva, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-1
Query Match          93.8%; Score 91; DB 9; Length 18;
Best Local Similarity 94.4%; Pred. No. 9.5e-08;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AALEAKICHQIEYFGDF 18
Db 1 AALEAKICHQIEYFGDF 18

RESULT 3
US-09-836-073-14
; Sequence 14, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baldeva, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bovine
US-09-836-073-14
Query Match          93.8%; Score 91; DB 9; Length 18;
Best Local Similarity 94.4%; Pred. No. 9.5e-08;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AALEAKICHQIEYFGDF 18
Db 1 AALEAKICHQIEYFGDF 18

RESULT 4
US-10-170-385-477
; Sequence 477, Application US/10170385
; Publication No. US20030203372A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingsman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 532682000100
; CURRENT APPLICATION NUMBER: US/10/170,385
```

```

; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 477
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-170-385-477
Query Match          93.8%; Score 91; DB 12; Length 408;
Best Local Similarity 94.4%; Pred. No. 2.4e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AALEAKICHQIEYFGDF 18
Db 11 AALEAKICHQIEYFGDF 28

RESULT 5
US-09-925-298-695
; Sequence 695, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 695
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-298-695
Query Match          93.8%; Score 91; DB 12; Length 460;
Best Local Similarity 94.4%; Pred. No. 2.7e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AALEAKICHQIEYFGDF 18
Db 63 AALEAKICHQIEYFGDF 80

RESULT 6
US-10-102-806-695
; Sequence 695, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 695
; LENGTH: 460
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:52:06 ; Search time 40.324 Seconds

(without alignments)
143.151 Million cell updates/sec

Title: US-09-836-073-12

Sequence: 1 AALEAKICHQIEYQGF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubppa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubppa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubppa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97	100.0	18	US-09-836-073-12	Sequence 12, Appl
2	91	93.8	18	US-09-836-073-1	Sequence 1, Appl
3	91	93.8	18	US-09-836-073-14	Sequence 14, Appl
4	91	93.8	408	US-10-170-385-477	Sequence 477, App
5	91	93.8	460	US-09-825-298-695	Sequence 695, App
6	91	93.8	460	US-10-102-806-695	Sequence 695, App
7	91	93.8	460	US-10-264-049-2643	Sequence 2643, App
8	89	91.8	18	US-09-836-073-8	Sequence 8, Appl
9	87	89.7	17	US-09-836-073-13	Sequence 13, Appl
10	85	87.6	18	US-09-836-073-9	Sequence 9, Appl
11	85	85.6	18	US-09-836-073-11	Sequence 11, Appl
12	82	84.5	18	US-09-836-073-10	Sequence 10, Appl
13	80.5	83.0	19	US-09-836-073-16	Sequence 16, Appl
14	79	81.4	18	US-09-836-073-2	Sequence 2, Appl
15	79	81.4	18	US-09-836-073-3	Sequence 3, Appl

15	79	81.4	18	9	US-09-836-073-4	Sequence 4, Appl
17	76	78.4	18	9	US-09-836-073-7	Sequence 7, Appl
18	74	76.3	18	9	US-09-836-073-6	Sequence 6, Appl
19	68	70.1	18	9	US-09-836-073-15	Sequence 15, Appl
20	67	69.1	18	9	US-09-836-073-5	Sequence 5, Appl
21	65	67.0	21	15	US-10-376-121A-20	Sequence 20, Appl
22	60	61.9	38	12	US-10-325-810-214	Sequence 214, App
23	51	52.6	38	9	US-09-843-676-25	Sequence 25, Appl
24	51	52.6	38	9	US-09-766-253-25	Sequence 25, Appl
25	51	52.6	38	10	US-09-438-486-25	Sequence 25, App
26	51	52.6	38	12	US-10-325-810-215	Sequence 215, App
27	51	52.6	38	12	US-10-053-758-25	Sequence 25, Appl
28	51	52.6	38	14	US-10-054-293-25	Sequence 25, Appl
29	51	52.6	38	14	US-10-054-611-25	Sequence 25, Appl
30	50.5	52.1	37	9	US-09-843-676-24	Sequence 24, Appl
31	50.5	52.1	37	9	US-09-766-253-24	Sequence 24, Appl
32	50.5	52.1	37	10	US-09-438-486-24	Sequence 24, Appl
33	50.5	52.1	37	14	US-10-053-758-24	Sequence 24, Appl
34	50.5	52.1	37	14	US-10-054-293-24	Sequence 24, Appl
35	50.5	52.1	37	14	US-10-054-611-24	Sequence 24, Appl
36	49	50.5	16	9	US-09-836-073-19	Sequence 19, Appl
37	48	49.5	182	12	US-10-282-132A-49328	Sequence 49328, A
38	48	49.5	696	16	US-10-437-963-200901	Sequence 200901, A
39	48	49.5	1527	16	US-10-437-963-200899	Sequence 200899, A
40	47	48.5	654	15	US-10-094-749-1854	Sequence 1854, App
41	47	48.5	658	15	US-10-210-130-66	Sequence 66, Appl
42	47	48.5	691	16	US-10-398-037-6	Sequence 6, Appl
43	46	47.4	343	16	US-10-437-963-111769	Sequence 111769, A
44	45	46.4	161	16	US-10-437-963-201436	Sequence 201436, A
45	45	46.4	224	15	US-10-094-749-2904	Sequence 2904, App

ALIGNMENTS

RESULT 1
US-09-836-073-12
; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Daesupda, Asim
; APPLICANT: Daes, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; TYPE: PRT
; LENGTH: 18
; ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match 100.0%; Score 97; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1e-08; 0;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYQGF 18
DB 1 AALEAKICHQIEYQGF 18

RESULT 2
US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Daesupda, Asim
; APPLICANT: Daes, S.

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-24

Query Match 52.1%; Score 50.5; DB 3; Length 37;
Best Local Similarity 83.3%; Pred. No. 0.065; 1; Indels 1; Gaps 1;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 7 ICHQIEYFGDF 18
||| |||
1 ICHQ-EYFGDF 11

Db 1 ICHQ-EYFGDF 11

RESULT 15
US-09-430-323-24
Sequence 24, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin
Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-430-323-24

Query Match 52.1%; Score 50.5; DB 4; Length 37;
Best Local Similarity 83.3%; Pred. No. 0.065; 1; Indels 1; Gaps 1;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 7 ICHQIEYFGDF 18
||| |||
1 ICHQ-EYFGDF 11

Db 1 ICHQ-EYFGDF 11

Search completed: September 10, 2004, 18:05:09
Job time : 16.8883 secs

FILED DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 215:
US-09-721-456-215

Query Match 52.6%; Score 51; DB 4; Length 38;
Best Local Similarity 81.8%; Pred. No. 0.055;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHQIEYFGD 17
DB 1 ICHQIEYFGD 11

RESULT 13
US-08-851-843A-24
Sequence 24, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017

FILED DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-24

Query Match 52.1%; Score 50.5; DB 3; Length 37;
Best Local Similarity 83.3%; Pred. No. 0.065;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICHQIEYFGDF 18
DB 1 ICHQIEYFGDF 11

RESULT 14
US-08-854-050-24
Sequence 24, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6261836el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536

FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-430-323-25

Query Match 52.6%; Score 51; DB 4; Length 38;
Best Local Similarity 81.8%; Pred. No. 0.055; 2; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 2;

Qy 7 ICHQIEYQFGD 17
Db 1 ICEQIEYFQGD 11

RESULT 11
US-09-402-181B-215
Sequence 215, Application US/09402181B
Patent No. 6610839
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 633
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,181B
FILING DATE: 29-Sep-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ausenhus, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
MOLECULE TYPE: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 215:
US-09-402-181B-215

Query Match 52.6%; Score 51; DB 4; Length 38;
Best Local Similarity 81.8%; Pred. No. 0.055; 2; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 2;

Qy 7 ICHQIEYQFGD 17
Db 1 ICEQIEYFQGD 11

RESULT 12
US-09-721-456-215
Sequence 215, Application US/09721456
Patent No. 6617110
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/721,456
FILING DATE: 22-No. 6617110-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017

FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-215

Query Match 52.6%; Score 51; DB 3; Length 38;
Best Local Similarity 81.8%; Pred. No. 0.055;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHQIEYFGD 17
DB 1 ICEQIEYFGD 11

RESULT 9
US-08-854-050-25
Sequence 25, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6261836el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-25

Query Match 52.6%; Score 51; DB 3; Length 38;
Best Local Similarity 81.8%; Pred. No. 0.055;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHQIEYFGD 17
DB 1 ICEQIEYFGD 11

RESULT 10
US-09-430-323-25
Sequence 25, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843

FILED DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 214:
US-09-721-456-214

Query Match 61.9%; Score 60; DB 4; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.0015;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHOEYFGDF 18
DB 1 ICHOEYFGDF 12

RESULT 7
US-08-851-843A-25
Sequence 25, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Hatley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809e1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-25

Query Match 52.6%; Score 51; DB 3; Length 38;
Best Local Similarity 81.8%; Pred. No. 0.055;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHOEYFGD 17
DB 1 ICHOEYFGD 11

RESULT 8
US-08-974-549A-215
Sequence 215, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Hatley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050

APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-874-549A-214

Query Match 61.9%; Score 60; DB 3; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.0015; 2; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 2;

OY 7 ICHQIEYFGDF 18
Db 1 ICHQKEYFGDF 12

RESULT 5
US-09-402-181B-214
Sequence 214, Application US/09402181B
Patent No. 6610839
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 633
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,181B
FILING DATE: 29-Sep-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503

FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ausubius, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 214:
US-09-402-181B-214

Query Match 61.9%; Score 60; DB 4; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.0015; 2; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 2;

OY 7 ICHQIEYFGDF 18
Db 1 ICHQKEYFGDF 12

RESULT 6
US-09-721-456-214
Sequence 214, Application US/09721456
Patent No. 6617110
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/721,456
FILING DATE: 22-No. 6617110-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312

PRIOR FILING DATE: 1997-10-06
PRIOR APPLICATION NUMBER: 08/321,427
PRIOR FILING DATE: 1994-10-11
PRIOR APPLICATION NUMBER: 60/086,527
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO: 4
LENGTH: 18
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: B-LAP
US-09-316-630-4

Query Match 93.8%; Score 91; DB 3; Length 18;
Best Local Similarity 94.4%; Pred. No. 3.2e-09;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AALEKICHOIEYFGDF 18
Db 1 AALEKICHOIEYFGDF 18

RESULT 3

US-08-475-955-20
Sequence 20, Application US/08475955
Patent No. 6641813
GENERAL INFORMATION:
APPLICANT: Harley, John
TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS OF
NUMBER OF SEQUENCES: 218
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center, 1250 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,955
FILING DATE: June 7, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/667,819
FILING DATE: April 13, 1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/648,205
FILING DATE: January 31, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/472,947
FILING DATE: January 31, 1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMF114C1P(2) DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) -873-8794
TELEFAX: (404) -873-8795
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Binding-site
LOCATION: 8..15
US-08-475-955-20

Query Match 67.0%; Score 65; DB 4; Length 21;
Best Local Similarity 91.7%; Pred. No. 0.00011;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ICHOIEYFGDF 18
Db 1 ICHOIEYFGDF 12

RESULT 4

US-08-974-549A-214
Sequence 214, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:43:01; Search time 15.8893 seconds
(without alignments)
58.488 Million cell updates/sec

Title: US-09-836-073-12

Perfect score: 97

Sequence: 1 AALEAKICHOIEYFGDF 18

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:
1: /cgnt2_6/prodata/2/iaa/5A_COMB.pep.*
2: /cgnt2_6/prodata/2/iaa/5B_COMB.pep.*
3: /cgnt2_6/prodata/2/iaa/6A_COMB.pep.*
4: /cgnt2_6/prodata/2/iaa/6B_COMB.pep.*
5: /cgnt2_6/prodata/2/iaa/PCTUS_COMB.pep.*
6: /cgnt2_6/prodata/2/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	93.8	18	US-09-316-630-3	Sequence 3, Appl1
2	91	93.8	18	US-09-316-630-4	Sequence 4, Appl1
3	65	67.0	21	US-08-475-955-20	Sequence 20, Appl1
4	60	61.9	38	US-08-974-549A-214	Sequence 214, App
5	60	61.9	38	US-09-402-181B-214	Sequence 214, App
6	60	61.9	38	US-09-721-456-214	Sequence 214, App
7	51	52.6	38	US-08-851-843A-25	Sequence 25, Appl
8	51	52.6	38	US-08-974-549A-215	Sequence 215, App
9	51	52.6	38	US-08-854-050-25	Sequence 25, Appl
10	51	52.6	38	US-09-430-323-25	Sequence 215, App
11	51	52.6	38	US-09-402-181B-215	Sequence 215, App
12	51	52.6	38	US-09-721-456-215	Sequence 215, App
13	50.5	52.1	37	US-08-851-843A-24	Sequence 24, Appl
14	50.5	52.1	37	US-08-854-050-24	Sequence 24, Appl
15	50.5	52.1	37	US-09-430-323-24	Sequence 24, Appl
16	44.5	45.9	124	US-09-489-039A-9103	Sequence 9103, App
17	44	45.4	431	US-08-311-023-2	Sequence 4, Appl1
18	42	43.3	456	US-08-205-719-4	Sequence 4, Appl1
19	42	43.3	456	US-08-431-517F-6	Sequence 6, Appl1
20	42	43.3	482	US-08-431-517F-5	Sequence 5, Appl1
21	42	43.3	482	5245013-2	Patent No. 5245013
22	41	42.3	699	US-09-543-881A-5118	Sequence 5118, App
23	40	41.2	39	US-08-851-843A-26	Sequence 26, Appl
24	40	41.2	39	US-08-974-549A-216	Sequence 216, App
25	40	41.2	39	US-08-854-050-26	Sequence 26, Appl
26	40	41.2	39	US-09-430-323-26	Sequence 26, Appl
27	40	41.2	39	US-09-402-181B-216	Sequence 216, App

28	40	41.2	39	4	US-09-721-456-216	Sequence 216, App
29	40	41.2	483	3	US-09-027-166-7	Sequence 7, Appl1
30	40	41.2	658	4	US-09-252-991A-25861	Sequence 25861, A
31	40	41.2	1068	1	US-08-537-210A-2	Sequence 2, Appl1
32	40	41.2	1068	1	US-09-113-825-2	Sequence 2, Appl1
33	40	41.2	2556	1	US-08-185-423-17	Sequence 17, Appl
34	40	41.2	2556	1	US-08-083-590A-20	Sequence 20, Appl
35	40	41.2	2556	3	US-08-532-384-20	Sequence 20, Appl
36	40	41.2	2556	1	US-08-899-232-2	Sequence 2, Appl1
37	39	40.2	158	2	US-08-933-750C-15	Sequence 15, Appl
38	39	40.2	158	3	US-09-234-613-15	Sequence 15, Appl
39	39	40.2	212	4	US-08-937-067-4	Sequence 4, Appl1
40	39	40.2	295	4	US-08-937-067-2	Sequence 2, Appl1
41	39	40.2	295	4	US-09-148-545-179	Sequence 179, App
42	39	40.2	296	4	US-09-148-545-237	Sequence 237, App
43	39	40.2	811	4	US-09-199-637A-93	Sequence 93, Appl
44	39	40.2	812	1	US-08-248-629A-1	Sequence 1, Appl1
45	39	40.2	812	1	US-08-451-932-1	Sequence 1, Appl1

ALIGNMENTS

```
RESULT 1
US-09-316-630-3
Sequence 3, Application US/09316630
Patent No. 6291637
GENERAL INFORMATION:
APPLICANT: Das, Saumitra
APPLICANT: Dasgupta, Asim
TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
TITLE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
FILE REFERENCE: 22000-20548-21
CURRENT APPLICATION NUMBER: US/09/316,630
PRIOR FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: 08/321,427
PRIOR FILING DATE: 1997-10-06
PRIOR APPLICATION NUMBER: 08/817,953
PRIOR FILING DATE: 1997-10-06
PRIOR APPLICATION NUMBER: 08/321,427
PRIOR FILING DATE: 1994-10-11
PRIOR APPLICATION NUMBER: 60/086,527
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 18
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3
Query Match 93.8%; Score 91; DB 3; Length 18;
Best Local Similarity 94.4%; Pred. No. 3.2e+09;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Cy 1 AALEAKICHOIEYFGDF 18
Db 1 AALEAKICHOIEYFGDF 18
RESULT 2
US-09-316-630-4
Sequence 4, Application US/09316630
Patent No. 6291637
GENERAL INFORMATION:
APPLICANT: Das, Saumitra
APPLICANT: Dasgupta, Asim
TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
TITLE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
FILE REFERENCE: 22000-20548-21
CURRENT APPLICATION NUMBER: US/09/316,630
CURRENT FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: 08/817,953
```

DR N-PSDB; AAS72604.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
PS Claim 20; SEQ ID NO 38776; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (II) is useful in gene therapy techniques to restore normal
CC activity of (I) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 439 AA;

Query Match 93.8%; Score 91; DB 4; Length 439;
Best Local Similarity 94.4%; Pred. No. 7.3e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYFGDF 18
| | | | | | | | | | | | | | | | | |
Db 41 AALEAKICHQIEYFGDF 58

RESULT 15

AAB58987 standard; protein; 460 AA.

AC AAB58987;

DT 27-MAR-2001 (first entry)

DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 695.

XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW neurologic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antineoplastic; antitumor; antiparasitic; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease.
XX

OS Homo sapiens.

PN WO200055173-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US005881.

PR 12-MAR-1999; 99US-0124270P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;
XX
XX WPI, 2000-611515/58.
DR N-PSDB; AAF21890.
XX
XX New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention, treatment
PT and diagnosis of cancer, immune disorders, cardiovascular disorders and
PT neurological diseases.
PS Claim 11; Page 1149-11150; 1299bp; English.
XX
XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive; neurologic;
CC neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
CC antineoplastic; antitumor; antiparasitic; anticonvulsant; antibacterial;
CC antifungal; antiparasitic and cardiac activity. The polynucleotide and
CC protein sequences are used in the diagnosis of cancer, particularly
CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
CC and agonists may also be used in the diagnosis, prevention and treatment
CC of immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC cardiovascular disorders such as myocardial ischaemia; wound healing;
CC neurological diseases such as cerebral anoxia and epilepsy; and
CC infectious diseases
XX
SQ Sequence 460 AA;

Query Match 93.8%; Score 91; DB 3; Length 460;
Best Local Similarity 94.4%; Pred. No. 7.7e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYFGDF 18
| | | | | | | | | | | | | | | | | |
Db 63 AALEAKICHQIEYFGDF 80

Search completed: September 10, 2004, 17:51:48
Job time : 58.4128 secs

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 415 AA;

Query Match 93.8%; Score 91; DB 7; Length 415;
 Best Local Similarity 94.4%; Pred. No. 6.9e-07;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AALEAKICHOIEYFGDF 18
 |||||
 DB 11 AALEAKICHOIEYFGDF 28

RESULT 13

ID ADE63993 standard; protein; 415 AA.

AC ADE63993;

DT 29-JAN-2004 (first entry)

DE Rat Protein P38656, SEQ ID NO 9939.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KM chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

PN WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

PA (GEHO) GEN HOSPITAL CORP.

PI (FARB) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M,

DR WPI; 2003-268312/26.

DR GENBANK; P38656.

PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017p; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 415 AA;

Query Match 93.8%; Score 91; DB 7; Length 415;
 Best Local Similarity 94.4%; Pred. No. 6.9e-07;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AALEAKICHOIEYFGDF 18
 |||||
 DB 11 AALEAKICHOIEYFGDF 28

RESULT 14

ID ABG08417 standard; protein; 439 AA.

AC ABG08417;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #8408.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT,

DR WPI; 2001-639362/73.

CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 408 AA;

Query Match 93.8%; Score 91; DB 7; Length 408;
Best Local Similarity 94.4%; Pred. No. 6.7e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AALEAKICHQIEYFGDF 18
Db 11 AALEAKICHQIEYFGDF 28
|||||

RESULT 11

ADE62859 ADE62859 standard; protein; 408 AA.

XX ADE62859;

DT 29-JAN-2004 (first entry)

DE Human Protein P05455, SEQ ID NO 8793.

XX Human; pain; neuronal tissue; gene therapy;

XX spinal segmental nerve injury; chronic constriction injury; CCI;

XX spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; P05455.

XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 408 AA;

Query Match 93.8%; Score 91; DB 7; Length 408;
Best Local Similarity 94.4%; Pred. No. 6.7e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AALEAKICHQIEYFGDF 18
Db 11 AALEAKICHQIEYFGDF 28
|||||

RESULT 12

ADE63989 ADE63989 standard; protein; 415 AA.

XX ADE63989;

DT 29-JAN-2004 (first entry)

DE Rat Protein P38656, SEQ ID NO 9935.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; P38656.

XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from MIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 408 AA;

Query Match 93.8%; Score 91; DB 7; Length 408;

Best Local Similarity 94.4%; Pred. No. 6.7e-07; Mismatches 1; Indels 0; Gaps 0;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 9
ADD6272
ID ADD6272 standard; protein; 408 AA.

ADD6272;

29-JAN-2004 (first entry)

Human Protein P05455, SEQ ID NO 11947.

Human; pain; neuronal tissue; gene therapy;

spinal segmental nerve injury; chronic constriction injury; CCI;

spared nerve injury; SNI; Chung.

Homo sapiens.

MO2003016475-A2.

27-FEB-2003.

14-AUG-2002; 2002WO-US025765.

14-AUG-2001; 2001US-0312147P.

01-NOV-2001; 2001US-0346382P.

26-NOV-2001; 2001US-033347P.

(GEHO) GEN HOSPITAL CORP.

(FARB) BAYER AG.

Woolf C, D'urso D, Befort K, Costigan M;

WPI; 2003-268312/26.

GENBANK; P05455.

Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the

polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from MIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 408 AA;

Query Match 93.8%; Score 91; DB 7; Length 408;

Best Local Similarity 94.4%; Pred. No. 6.7e-07; Mismatches 1; Indels 0; Gaps 0;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 10
ADE63995
ID ADE63995 standard; protein; 408 AA.

ADE63995;

29-JAN-2004 (first entry)

Human Protein P05455, SEQ ID NO 9941.

Human; pain; neuronal tissue; gene therapy;

spinal segmental nerve injury; chronic constriction injury; CCI;

spared nerve injury; SNI; Chung.

Homo sapiens.

MO2003016475-A2.

27-FEB-2003.

14-AUG-2002; 2002WO-US025765.

14-AUG-2001; 2001US-0312147P.

01-NOV-2001; 2001US-0346382P.

26-NOV-2001; 2001US-033347P.

(GEHO) GEN HOSPITAL CORP.

(FARB) BAYER AG.

Woolf C, D'urso D, Befort K, Costigan M;

WPI; 2003-268312/26.

GENBANK; P05455.

Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence

CC is clinically related to the Ro protein that is highly common among
CC autoimmune patients. La and Ro antigens sometimes reside on the same
CC cellular ribonucleoprotein particle; most La patients contain some Ro
CC antibodies and vice versa. La cDNA has been isolated from a human liver
CC library. (Updated on 25-MAR-2003 to correct PF field.)
XX

SQL Sequence 408 AA;

Query Match 93.8%; Score 91; DB 2; Length 408;
Best Local Similarity 94.4%; Pred. No. 6.7e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AALEAKICHQIEYFGDF 18
|||
DB 11 AALEAKICHQIEYFGDF 28

RESULT 7

ID ABP65252 standard; protein; 408 AA.

XX ABP65252;

DT 12-NOV-2002 (first entry)

XX Hypoxia-regulated protein #126.

XX Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;
XX antiinflammatory; vulnary; gynecological; ophthalmological; vaccine;
XX hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;
XX ischaemic condition; reperfusion injury; retinopathy; neonatal stress;
XX preclapmsia; atherosclerosis; inflammatory condition; wound healing;
XX inflammation; erythropoiesis; hair loss; human.

XX Homo sapiens.

XX WO200246465-A2.

XX 13-JUN-2002.

XX 10-DEC-2001; 2001WO-GB005458.

XX 08-DEC-2000; 2000GB-00030076.

XX 08-FEB-2001; 2001GB-00003156.

XX 25-OCT-2001; 2001GB-00025666.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX White J, Mundy CR, Ward NR, Krigge D, Kingsman SM, Harris RA;

XX Rayner WN;

XX WPI; 2002-627238/67.

PT Identifying a gene involved in disease for treating hypoxia-regulated
PT conditions, comprises comparing the transcriptome/proteome of two cell
PT types under different conditions and identifying a differentially
PT regulated gene.

XX Claim 35; Page 425; 538pp; English.

CC The present invention relates to methods for identifying genes and
CC proteins that are implicated in a specific disease or physiological
CC condition. The method comprises comparing the transcriptome/proteome of a
CC specialised cell type implicated in a disease or condition with that of a
CC second specialised cell type, under two experimental conditions, and
CC identifying a gene that is differentially regulated in the two
CC specialised cell types under experimental conditions. ABV7873-ABV7816
CC and ABP65061-ABP65257 were identified using the methods of the invention.
CC The coding sequences and proteins are useful for treating a disease in a
CC patient, for manufacture of a medicament for treating hypoxia-regulated
CC conditions, and for regulating tumorigenesis, angiogenesis, apoptosis,
CC biological response to hypoxia conditions, or hypoxia-associated
CC pathology in a patient. The coding sequences and proteins are also useful

CC for monitoring the therapeutic treatment of a disease or physiological
CC condition, such as cancer, ischaemic conditions, reperfusion injury,
CC retinopathy, neonatal stress, preclapmsia, atherosclerosis, inflammatory
CC conditions, wound healing, inflammation, erythropoiesis or hair loss
XX

SQL Sequence 408 AA;

Query Match 93.8%; Score 91; DB 5; Length 408;
Best Local Similarity 94.4%; Pred. No. 6.7e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AALEAKICHQIEYFGDF 18
|||
DB 11 AALEAKICHQIEYFGDF 28

RESULT 8

ID ADE63991 standard; protein; 408 AA.

XX ADE63991;

DT 29-JAN-2004 (first entry)

XX Human Protein P05455, SEQ ID NO. 9937.

XX Human; pain; neuronal tissue; gene therapy;
XX spinal segmental nerve injury; chronic constriction injury; CCI;
XX spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GHEO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; P05455.

XX Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more

KM internal ribosome entry site initiated translation; Ia antigen protein;
 KM viral life cycle; antiviral therapy; pharmaceutical; veterinary;
 KM agricultural; horticultural; virucide; bovine.
 XX Bovinae.
 XX WO200283858-A2.
 XX PD 24-OCT-2002.
 XX PF 12-APR-2002; 2002WO-US011589.
 XX PR 16-APR-2001; 2001US-00836073.
 XX PA (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.
 XX PI Dasgupta A, Das S, Balda N;
 XX DR WPI; 2003-058634/05.
 XX PT New compound containing acidic and aromatic amino acids, useful as
 PT antiviral therapy in pharmaceutical, veterinary or
 PT agricultural/horticultural applications.
 XX PS Disclosure; Page 6; 19pp; English.
 XX CC The present invention relates to peptides and methods of inhibiting the
 CC replication of viruses that utilize internal ribosome entry site (IRES)
 CC initiated translation, and/or inhibiting viruses that utilize the Ia
 CC antigen protein (IAP) in any phase of their life cycle. The peptides of
 CC the invention compete with IAP and inhibit the utilisation of various
 CC biochemical and physiological functions of IAP required for a productive
 CC life cycle. The methods and compositions are useful as antiviral therapy
 CC in pharmaceutical, veterinary or agricultural/horticultural applications.
 CC ABG72101-ABG72119 represent peptides useful as antiviral agents
 XX SQ Sequence 18 AA;
 XX
 XX Query Match 93.8%; Score 91; DB 6; Length 18;
 XX Best Local Similarity 94.4%; Pred. No. 2.2e-08;
 XX Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AALEAKICHOIEYQFGDF 18
 DB 1 AALEAKICHOIEYFGDF 18
 XX
 XX RESULT 5
 XX ID AAG01351 standard; protein; 92 AA.
 XX AC AAG01351;
 XX DT 06-OCT-2000 (first entry)
 XX DE Human secreted protein, SEQ ID NO: 5432.
 XX KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KM gene therapy; chromosome mapping.
 XX OS Homo sapiens.
 XX PN BP1033401-A2.
 XX PD 06-SEP-2000.
 XX PF 21-FEB-2000; 2000EP-00200610.
 XX PR 26-FEB-1999; 99US-0122487P.
 XX PA (GEST) GENSET.
 XX PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX DR WPI; 2000-500381/45.
 XX DR N-PSDB; AAC01357.
 XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
 XX PS Claim 13; SEQ ID NO 5432; 71pp + Sequence listing; English.
 XX CC The present sequence is a polypeptide encoded by one of a large number of
 CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
 CC prepared from total human RNAs or polyA+ RNAs derived from 30 different
 CC tissues. EST sequences usually correspond mainly to the 3' untranslated
 CC region (UTR) of the mRNA because they are often obtained for oligo-dT
 CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
 CC sequences derived from the 5' ends of mRNAs and even in those cases where
 CC longer cDNA sequences have been obtained, the full 5' UTR is rarely
 CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can
 CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
 CC are also used in diagnostic, forensic, gene therapy and chromosome
 CC mapping procedures. They are used to obtain upstream regulatory sequences
 CC and to design expression and secretion vectors
 XX SQ Sequence 92 AA;
 XX
 XX Query Match 93.8%; Score 91; DB 3; Length 92;
 XX Best Local Similarity 94.4%; Pred. No. 1.3e-07;
 XX Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AALEAKICHOIEYQFGDF 18
 DB 11 AALEAKICHOIEYFGDF 28
 XX
 XX RESULT 6
 XX ID AAW03716 standard; protein; 408 AA.
 XX AC AAW03716;
 XX DT 25-MAR-2003 (revised)
 DT 12-MAR-1997 (first entry)
 XX DE Human autoantigen La(SS-B).
 XX KM Autoimmune disease; Ia autoantigen; Sjogren's syndrome;
 KM systemic lupus erythematosus; diagnosis.
 XX OS Homo sapiens.
 XX PN US5541291-A.
 XX PD 30-JUL-1996.
 XX PF 27-MAY-1987; 87US-00054871.
 XX PR 31-DEC-1984; 84US-00687908.
 XX PA (UYDU-) UNIV DUKE.
 XX PI Keene JD;
 XX DR WPI; 1996-362015/36.
 XX PT Auto-antigen U2-RNP, associated with severe polymyositis-scleroderma
 PT overlap syndrome - useful for diagnosis and treatment of autoimmune
 PT diseases.
 XX PS Disclosure; Col 15-16; 21pp; English.
 XX CC The human lupus antigen (Ia) is diagnostic for Sjogren's syndrome, as
 CC well as occurring in systemic lupus erythematosus patients. The Ia protein

Query Match 100.0%; Score 97; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 2.1e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AALEAKICHQIEYFGDF 18
 DB 1 AALEAKICHQIEYFGDF 18

RESULT 2
 ID AAY52200 standard; peptide; 18 AA.

XX AAY52200;
 AC AAY52200;
 XX 14-MAR-2000 (first entry)
 DT 14-MAR-2000 (first entry)
 XX Human Ia autoantigen peptide (LAP).

DE La autoantigen; LAP, internal ribosome entry site, IRES, translation;
 KM viral replication; IRNA; antiviral agent; picornavirus; flavivirus;
 KM coronavirus; hepatitis virus; rhadovirus; adenovirus; coxsackie virus;
 KM parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
 KM foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
 KM vesicular stomatitis virus.

OS Homo sapiens.
 XX MO3961613-A2.
 XX 02-DEC-1999.
 PD 21-MAY-1999; 99WO-US011281.
 XX 22-MAY-1998; 98US-0086527P.
 PR (RESC) UNIV CALIFORNIA.
 XX Das S, Dasgupta A;
 PI WPI; 2000-062712/05.

DR WPI; 2000-062712/05.
 XX New yeast inhibitory peptide useful for inhibiting viral protein
 PT translation and replication.
 PS Claim 5; Page 57; 81pp; English.

CC This sequence is the Ia autoantigen binding domain (LAP). LAP is a
 CC cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 CC translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors from
 CC binding at the site. The peptide can be used alone or in combination with
 CC an inhibitor RNA (IRNA see AAY5200). The LAP peptide is useful as an
 CC antiviral agent, which works through the inhibition of mRNA translation,
 CC especially viral mRNA. Examples of viruses which can be inhibited are
 CC picornavirus, flavivirus, coronavirus, hepatitis A B or C viruses,
 CC rhadovirus, adenovirus, and parainfluenza virus, poliovirus, rhinovirus,
 CC coxsackie virus, encephalomyocarditis virus, foot-and-mouth disease
 CC virus, echo virus, hepatitis C virus, infectious bronchitis virus, duck
 CC and human hepatitis B virus, and vesicular stomatitis virus. The peptide
 CC also inhibits replication of the above viruses. The LAP peptide
 CC selectively inhibits viral protein translation, and is therefore not
 CC toxic to the host cell

CC Sequence 18 AA;

Query Match 93.8%; Score 91; DB 3; Length 18;
 Best Local Similarity 94.4%; Pred. No. 2.2e-08;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AALEAKICHQIEYFGDF 18

DB 1 AALEAKICHQIEYFGDF 18

RESULT 3
 ID ABG72101 standard; peptide; 18 AA.

XX ABG72101;
 AC ABG72101;
 XX 28-JAN-2003 (first entry)

DE Viral replication inhibiting peptide, LAP.

KM Viral replication inhibitor; IRES initiated translation; LAP;
 KM internal ribosome entry site initiated translation; Ia antigen protein;
 KM viral life cycle; antiviral therapy; pharmaceutical; veterinary;
 KM agricultural; horticultural; virucide.

OS Unidentified.
 XX WO200283858-A2.
 XX 24-OCT-2002.

PD 12-APR-2002; 2002WO-US011589.
 PF 16-APR-2001; 2001US-00836073.

PR (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.
 XX Dasgupta A, Das S, Baidya N;
 PI WPI; 2003-058634/05.

DR New compound containing acidic and aromatic amino acids, useful as
 PT antiviral therapy in pharmaceutical, veterinary or
 PT agricultural/horticultural applications.
 PS Claim 1; Page 15; 19pp; English.

CC The present invention relates to peptides and methods of inhibiting the
 CC replication of viruses that utilize internal ribosome entry site (IRES)
 CC initiated translation, and/or inhibiting viruses that utilize the Ia
 CC antigen protein (LAP) in any phase of their life cycle. The peptides of
 CC the invention compete with LAP and inhibit the utilization of various
 CC biochemical and physiological functions of LAP required for a productive
 CC life cycle. The methods and compositions are useful as antiviral therapy
 CC in pharmaceutical, veterinary or agricultural/horticultural applications.
 CC ABG72101-ABG72119 represent peptide useful as antiviral agents

CC Sequence 18 AA;

Query Match 93.8%; Score 91; DB 6; Length 18;
 Best Local Similarity 94.4%; Pred. No. 2.2e-08;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AALEAKICHQIEYFGDF 18
 DB 1 AALEAKICHQIEYFGDF 18

RESULT 4
 ID ABG72114 standard; peptide; 18 AA.

XX ABG72114;
 AC ABG72114;

XX 28-JAN-2003 (first entry)
 DT Viral replication inhibiting peptide, BOVINE.
 XX Viral replication inhibitor; IRES initiated translation; LAP;

KM Viral replication inhibitor; IRES initiated translation; LAP;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 16:58:20 ; Search time 56.3129 Seconds
(without alignments)
90.314 Million cell updates/sec

Title: US-09-836-073-12

Perfect score: 97

Sequence: 1 AALEAKICHQIEYGFQDF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	100.0	18	6	ABG72112
2	91	93.8	18	3	AAI52200
3	91	93.8	18	6	ABG72101
4	91	93.8	18	6	ABG72114
5	91	93.8	92	3	AAAG01351
6	91	93.8	408	2	AAW03716
7	91	93.8	408	5	ABP65252
8	91	93.8	408	7	ADDE63991
9	91	93.8	408	7	ADDA6272
10	91	93.8	408	7	ADDE63995
11	91	93.8	408	7	ADDE62859
12	91	93.8	415	7	ADDE63989
13	91	93.8	415	7	ADDE63993
14	91	93.8	439	4	ABG08417
15	91	93.8	460	3	AAAB58987
16	91	93.8	460	5	ABP41511
17	89	91.8	18	6	ABG72113
18	87	89.7	17	6	ABG72110
19	85	87.6	18	6	ABG72109
20	83	85.6	18	6	ABG72111
21	82	84.5	18	6	ABG72110
22	80.5	83.0	19	6	ABG72116
23	79	81.4	18	6	ABG72102
24	79	81.4	18	6	ABG72103
25	79	81.4	18	6	ABG72104

26	76	78.4	18	6	ABG72107
27	74	76.3	18	6	ABG72106
28	68	70.1	18	6	ABG72115
29	67	69.1	18	6	ABG72105
30	65	67.0	21	2	AAAR43394
31	49	50.5	16	6	ABG72119
32	48	49.5	182	6	ABU21404
33	47	48.5	654	6	ADA54286
34	47	48.5	658	7	ADBA47704
35	47	48.5	691	5	ABP43482
36	46	47.4	517	5	AAU74627
37	46	47.4	519	5	ABU71162
38	46	47.4	519	5	AAU74628
39	45	46.4	135	4	ABG27058
40	45	46.4	224	5	ABB89645
41	45	46.4	224	6	ADA55336
42	44	45.4	129	6	ADA55269
43	44	45.4	160	4	AAAG93460
44	44	45.4	203	4	AAW40110
45	44	45.4	204	3	AAAB59035

ALIGNMENTS

RESULT 1
ABG72112
ID ABG72112 standard; peptide; 18 AA.

XX ABG72112;

XX 28-JAN-2003 (first entry)

DE Viral replication inhibiting peptide, 633.

XX Viral replication inhibitor; IRES initiated translation; LAP;

XX internal ribosome entry site initiated translation; La antigen protein;

KW viral life cycle; antiviral therapy; pharmaceutical; veterinary;

KW agricultural; horticultural; virucide.

XX Unidentified.

XX WO200283858-A2.

XX 24-OCT-2002.

XX 12-APR-2002; 2002WO-US011589.

XX 16-APR-2001; 2001US-00836073.

XX (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.

XX Dasgupta A, Das S, Baidya N;

XX WPI; 2003-058634/05.

XX New compound containing acidic and aromatic amino acids, useful as

XX antiviral therapy in pharmaceutical, veterinary or

XX agricultural/horticultural applications.

XX Claim 10; Page 16; 19pp; English.

XX The present invention relates to peptides and methods of inhibiting the

XX replication of viruses that utilize internal ribosome entry site (IRES)

XX initiated translation, and/or inhibiting viruses that utilize the La

XX antigen protein (LAP) in any phase of their life cycle. The peptides of

XX biochemical and physiological functions of LAP required for a productive

XX life cycle. The methods and compositions are useful as antiviral therapy

XX in pharmaceutical, veterinary or agricultural/horticultural applications.

XX ABG72101-ABG72119 represent peptides useful as antiviral agents

XX Sequence 18 AA;

This Page Blank (uspto)

Mon Sep 13 09:36:23 2004

us-09-836-073-4.rapb

Search completed: September 10, 2004, 18:11:55
Job time : 40.424 secs


```
FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-11

Query Match      77.5%; Score 79; DB 9; Length 18;
Best Local Similarity 93.3%; Pred. No. 1.9e-05;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4 EAKICHQIEYFGDF 18
        |||||
        4 EAKICHQIEYFGDF 18

Db      4 EAKICHQIEYFGDF 18

RESULT 12
US-09-836-073-12
; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balgva, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match      77.5%; Score 79; DB 9; Length 18;
Best Local Similarity 93.3%; Pred. No. 1.9e-05;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4 EAKICHQIEYFGDF 18
        |||||
        4 EAKICHQIEYFGDF 18

Db      4 EAKICHQIEYFGDF 18

RESULT 13
US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balgva, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
```

```
US-09-836-073-10

Query Match      76.5%; Score 78; DB 9; Length 18;
Best Local Similarity 93.3%; Pred. No. 2.7e-05;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4 EAKICHQIEYFGDF 18
        |||||
        4 EAKICHQIEYFGDF 18

Db      4 EAKICHQIEYFGDF 18

RESULT 14
US-09-836-073-16
; Sequence 16, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balgva, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Rat
US-09-836-073-16

Query Match      75.0%; Score 76.5; DB 9; Length 19;
Best Local Similarity 93.8%; Pred. No. 4.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY      4 EAKICHQI-EYFGDF 18
        |||||
        4 EAKICHQIEYFGDF 19

Db      4 EAKICHQIEYFGDF 19

RESULT 15
US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balgva, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match      73.5%; Score 75; DB 9; Length 18;
Best Local Similarity 86.7%; Pred. No. 7.9e-05;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 EAKICHQIEYFGDF 18
        ||:|
        4 EAKICHQIEYFGDF 18

Db      4 EAKICHQIEYFGDF 18
```

US-10-170-385-477

Query Match 85.3%; Score 87; DB 12; Length 408;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
14 EAKICHOIEYFGDF 28

RESULT 7

US-09-925-298-695
; Sequence 695, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 695
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-298-695

Query Match 85.3%; Score 87; DB 12; Length 460;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
66 EAKICHOIEYFGDF 80

RESULT 8

US-10-102-806-695
; Sequence 695, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 695
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-806-695

Query Match 85.3%; Score 87; DB 14; Length 460;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
66 EAKICHOIEYFGDF 80

RESULT 9

US-10-264-049-2643
; Sequence 2643, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: Patentln Ver. 3.1
; SEQ ID NO 2643
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-2643

Query Match 85.3%; Score 87; DB 15; Length 460;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
66 EAKICHOIEYFGDF 80

RESULT 10

US-09-836-073-9
; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-9

Query Match 79.4%; Score 81; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.1e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGD 17
Db 4 EAKICHOIEYFGD 17

RESULT 11

US-09-836-073-11
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

```

; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-5

Query Match      88.2%; Score 90; DB 9; Length 18;
Best Local Similarity 88.9%; Pred. No. 3.5e-07;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 000EAKICHOIEYFGDF 18
Db 1 000EORQCHOIEYFGDF 18

RESULT 3
US-09-836-073-13
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Mouse
US-09-836-073-13

Query Match      85.3%; Score 87; DB 9; Length 17;
Best Local Similarity 100.0%; Pred. No. 9.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EAKICHOIEYFGDF 18
Db 3 EAKICHOIEYFGDF 17

RESULT 4
US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 16
```

```

; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-1

Query Match      85.3%; Score 87; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EAKICHOIEYFGDF 18
Db 4 EAKICHOIEYFGDF 18

RESULT 5
US-09-836-073-14
; Sequence 14, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bovine
US-09-836-073-14

Query Match      85.3%; Score 87; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EAKICHOIEYFGDF 18
Db 4 EAKICHOIEYFGDF 18

RESULT 6
US-10-170-385-477
; Sequence 477, Application US/10170385
; Publication No. US20030203372A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingsman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 532682000100
; CURRENT APPLICATION NUMBER: US/10/170,385
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 477
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Homo Sapiens
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

SW protein - protein search, using sw model

Run on: September 10, 2004, 17:52:06 ; Search time 40.324 Seconds
(without alignments)
143.151 Million cell updates/sec

Title: US-09-836-073-4

Perfect score: 102
Sequence: 1 QOGEAKICHQIEYFGDF 18

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
- 4: /cgn2_6/ptodata/1/pubppaa/US07_PUB.pep:*
- 5: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
- 6: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
- 7: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
- 18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	100.0	18	US-09-836-073-4	Sequence 4, Appli
2	90	88.2	18	US-09-836-073-5	Sequence 5, Appli
3	87	85.3	17	US-09-836-073-13	Sequence 13, Appli
4	87	85.3	18	US-09-836-073-1	Sequence 1, Appli
5	87	85.3	18	US-09-836-073-14	Sequence 14, Appli
6	87	85.3	408	US-10-170-385-477	Sequence 477, App
7	87	85.3	460	US-09-925-298-695	Sequence 695, App
8	87	85.3	460	US-10-102-806-695	Sequence 695, App
9	87	85.3	460	US-10-264-049-2643	Sequence 2643, Ap
10	81	79.4	18	US-09-836-073-9	Sequence 9, Appli
11	79	77.5	18	US-09-836-073-11	Sequence 11, Appli
12	79	77.5	18	US-09-836-073-12	Sequence 12, Appli
13	78	76.5	18	US-09-836-073-10	Sequence 10, Appli
14	76.5	75.0	19	US-09-836-073-16	Sequence 16, Appli
15	75	73.5	18	US-09-836-073-2	Sequence 2, Appli

16	75	73.5	18	US-09-836-073-3	Sequence 3, Appli
17	73	71.6	21	US-10-376-121A-20	Sequence 20, Appli
18	72	70.6	18	US-09-836-073-7	Sequence 7, Appli
19	72	70.6	18	US-09-836-073-15	Sequence 15, Appli
20	71	69.6	18	US-09-836-073-8	Sequence 8, Appli
21	68	66.7	38	US-10-325-810-214	Sequence 214, App
22	62	60.8	16	US-09-836-073-19	Sequence 19, Appli
23	59	57.8	38	US-09-843-676-25	Sequence 25, Appli
24	59	57.8	38	US-09-766-253-25	Sequence 25, Appli
25	59	57.8	38	US-09-438-486-25	Sequence 25, Appli
26	59	57.8	38	US-10-325-810-215	Sequence 215, App
27	59	57.8	38	US-10-053-758-25	Sequence 25, Appli
28	59	57.8	38	US-10-054-295-25	Sequence 25, Appli
29	59	57.8	38	US-10-054-611-25	Sequence 25, Appli
30	58.5	57.4	37	US-09-843-676-24	Sequence 24, Appli
31	58.5	57.4	37	US-09-766-253-24	Sequence 24, Appli
32	58.5	57.4	37	US-10-053-758-24	Sequence 24, Appli
33	58.5	57.4	37	US-10-054-295-24	Sequence 24, Appli
34	58.5	57.4	37	US-10-054-611-24	Sequence 24, Appli
35	58.5	57.4	37	US-10-054-611-24	Sequence 24, Appli
36	56	54.9	18	US-09-836-073-6	Sequence 6, Appli
37	51	50.0	190	US-10-437-963-184914	Sequence 184914, Sequence 272690,
38	49	48.0	488	US-10-424-599-272690	Sequence 125036,
39	49	48.0	922	US-10-437-963-125036	Sequence 26, Appli
40	48	47.1	39	US-09-843-676-26	Sequence 26, Appli
41	48	47.1	39	US-09-766-253-26	Sequence 26, Appli
42	48	47.1	39	US-09-438-486-26	Sequence 26, Appli
43	48	47.1	39	US-10-325-810-216	Sequence 216, App
44	48	47.1	39	US-10-053-758-26	Sequence 26, Appli
45	48	47.1	39	US-10-054-295-26	Sequence 26, Appli

ALIGNMENTS

RESULT 1
US-09-836-073-4
; Sequence 4, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Daegupta, Asim
; APPLICANT: Dae, S.
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836, 073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316, 630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-4

Query Match 100.0%; Score 102; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.7e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 QOGEAKICHQIEYFGDF 18
Db 1 QOGEAKICHQIEYFGDF 18
RESULT 2
US-09-836-073-5
; Sequence 5, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Daegupta, Asim
; APPLICANT: Dae, S.

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULAR TYPE: peptide
US-08-854-050-24

Query Match 57.4%; Score 58.5; DB 3; Length 37;
Best Local Similarity 91.7%; Pred. No. 0.008; 0; Indels 1; Gaps 1;
Matches 11; Conservative 0; Mismatches 0;

Qy 7 ICHQIEYFGDF 18
Db 1 ICHQ-EYFGDF 11

RESULT 15
US-09-430-323-24
Sequence 24, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin
Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULAR TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-430-323-24

Query Match 57.4%; Score 58.5; DB 4; Length 37;
Best Local Similarity 91.7%; Pred. No. 0.008; 0; Indels 1; Gaps 1;
Matches 11; Conservative 0; Mismatches 0;

Qy 7 ICHQIEYFGDF 18
Db 1 ICHQ-EYFGDF 11

Search completed: September 10, 2004, 18:05:08
Job time: 16.8883 secs

FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 215:
US-09-721-456-215

Query Match 57.8%; Score 59; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0069; 1; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 1;

QY 7 ICHQIEYFGD 17
DB 1 ICEQIEYFGD 11

RESULT 13
US-08-851-843A-24
Sequence 24, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809e1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-24

Query Match 57.4%; Score 58.5; DB 3; Length 37;
Best Local Similarity 91.7%; Pred. No. 0.008; 0; Indels 1; Gaps 1;
Matches 11; Conservative 0; Mismatches 0;

QY 7 ICHQIEYFGDF 18
DB 1 ICHQ-EYFGDF 11

RESULT 14
US-08-854-050-24
Sequence 24, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6261836e1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536

FILED DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-430-323-25

Query Match 57.8%; Score 59; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0069; 1; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 1;

QY 7 ICHQIEYYFGD 17
Db 1 ICEQIEYYFGD 11

RESULT 11
US-09-402-181B-215
Sequence 215, Application US/09402181B
Patent No. 6610839
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 613
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,181B
FILING DATE: 29-Sep-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050

FILED DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ausenhus, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0300
TELEFAX: (415) 576-0200
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 215:
US-09-402-181B-215

Query Match 57.8%; Score 59; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0069; 1; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 1;

QY 7 ICHQIEYYFGD 17
Db 1 ICEQIEYYFGD 11

RESULT 12
US-09-721-456-215
Sequence 215, Application US/09721456
Patent No. 6617110
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/721,456
FILING DATE: 22-NOV-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017

FILED DATE: 09-MAY-1997
PRIOR APPLICATION DATA: 536
CLASSIFICATION: 536
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-215

Query Match 57.8%; Score 59; DB 3; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0069;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIEYFGD 17
DB 1 ICEQIEYFGD 11

RESULT 9
US-08-854-050-25
Sequence 25, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cecch, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6261836el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843

FILED DATE: 06-MAY-1997
CLASSIFICATION: 536
APPLICATION NUMBER: 536
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-25

Query Match 57.8%; Score 59; DB 3; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0069;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIEYFGD 17
DB 1 ICEQIEYFGD 11

RESULT 10
US-09-430-323-25
Sequence 25, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cecch, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843

FILED DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 214:
US-09-721-456-214

Query Match 66.7%; Score 68; DB 4; Length 38;
Best Local Similarity 91.7%; Pred. No. 0.00024;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIEYFYFGDF 18
DB 1 ICHQIEYFYFGDF 12

RESULT 7
US-08-851-843A-25
Sequence 25, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin H.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-25

Query Match 57.8%; Score 59; DB 3; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0069;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIEYFYFGD 17
DB 1 ICEQIEYFYFGD 11

RESULT 8
US-08-974-549A-215
Sequence 215, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin H.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050

APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-214

Query Match 66.7%; Score 68; DB 3; Length 38;
Best Local Similarity 91.7%; Pred. No. 0.00024;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIEYFGDF 18
Db 1 ICHQIEYFGDF 12

RESULT 5
US-09-402-181B-214
Sequence 214, Application US/09402181B
Patent No. 6610839
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 633
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,181B
FILING DATE: 29-Sep-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503

FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Auesenhus, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 214:
US-09-402-181B-214

Query Match 66.7%; Score 68; DB 4; Length 38;
Best Local Similarity 91.7%; Pred. No. 0.00024;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIEYFGDF 18
Db 1 ICHQIEYFGDF 12

RESULT 6
US-09-721-456-214
Sequence 214, Application US/09721456
Patent No. 6617110
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/721,456
FILING DATE: 22-No. 6617110-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312

PRIOR FILING DATE: 1997-10-06
PRIOR APPLICATION NUMBER: 08/321,427
PRIOR FILING DATE: 1994-10-11
PRIOR APPLICATION NUMBER: 60/086,527
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 4
LENGTH: 18
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: B-LAP
US-09-316-630-4

Query Match 85.3%; Score 87; DB 3; Length 18;
Best Local Similarity 100.0%; Pred. No. 8.7e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHQIEYFGDF 18
DB 4 EAKICHQIEYFGDF 18

RESULT 3
US-08-475-955-20
Sequence 20, Application US/08475955
Patent No. 6641813
GENERAL INFORMATION:
APPLICANT: Harley, John
TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS OF
NUMBER OF SEQUENCES: 218
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabat
STREET: 2800 One Atlantic Center, 1250 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,955
FILING DATE: June 7, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/867,819
FILING DATE: April 13, 1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/648,205
FILING DATE: January 31, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/472,947
FILING DATE: January 31, 1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabat, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRP14CIP(2) DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-873-8795
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Binding-site
LOCATION: 8.15
US-08-475-955-20

Query Match 71.6%; Score 73; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICHQIEYFGDF 18
DB 1 ICHQIEYFGDF 12

RESULT 4
US-08-974-549A-214
Sequence 214, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:43:01 ; Search time 15.8883 Seconds
(without alignments)
58,488 Million cell updates/sec

Title: US-09-836-073-4

Perfect score: 102
Sequence: 1 QQQEAKICHOIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pep:*
- 2: /cgn2_6/ptodata/2/iaa/5B COMB.pep:*
- 3: /cgn2_6/ptodata/2/iaa/6A COMB.pep:*
- 4: /cgn2_6/ptodata/2/iaa/6B COMB.pep:*
- 5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pep:*
- 6: /cgn2_6/ptodata/2/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	85.3	18	3	US-09-316-630-3
2	87	85.3	18	3	US-09-316-630-4
3	73	71.6	21	4	US-08-475-955-20
4	68	66.7	38	3	US-08-974-549A-214
5	68	66.7	38	4	US-09-402-181B-214
6	68	66.7	38	4	US-09-721-456-214
7	59	57.8	38	3	US-08-851-843A-25
8	59	57.8	38	3	US-08-974-549A-215
9	59	57.8	38	3	US-08-854-050-25
10	59	57.8	38	4	US-09-430-323-25
11	59	57.8	38	4	US-09-402-181B-215
12	59	57.8	38	4	US-09-721-456-215
13	58.5	57.4	37	3	US-08-851-843A-24
14	58.5	57.4	37	3	US-08-854-050-24
15	58.5	57.4	37	3	US-09-430-323-24
16	48	47.1	39	3	US-08-851-843A-26
17	48	47.1	39	3	US-08-974-549A-216
18	48	47.1	39	3	US-08-854-050-26
19	48	47.1	39	4	US-09-430-323-26
20	48	47.1	39	4	US-09-402-181B-216
21	48	47.1	39	4	US-09-721-456-216
22	41	40.2	497	4	US-09-489-039A-8926
23	41	40.2	1220	2	US-08-680-326-38
24	40	39.2	69	4	US-09-621-976-7385
25	40	39.2	124	4	US-09-489-039A-9103
26	40	39.2	431	1	US-08-311-023-2
27	40	39.2	582	4	US-09-976-594-733

28	39	38.2	122	4	US-09-732-210-62	Sequence 62, Appl
29	39	38.2	246	4	US-09-104-678A-2	Sequence 2, Appl
30	39	38.2	247	4	US-09-548-938A-12	Sequence 12, Appl
31	39	38.2	333	4	US-09-252-991A-31345	Sequence 31345, A
32	39	38.2	467	4	US-09-252-991A-24329	Sequence 24329, A
33	39	38.2	463	3	US-09-027-166-7	Sequence 7, Appl
34	39	38.2	754	2	US-08-941-262-1	Sequence 1, Appl
35	39	38.2	755	2	US-08-941-262-3	Sequence 1, Appl
36	39	38.2	817	4	US-09-252-991A-27609	Sequence 27609, A
37	38.5	37.7	412	4	US-09-543-681A-4298	Sequence 4298, Ap
38	38	37.3	141	4	US-09-540-236-2332	Sequence 2332, Ap
39	38	37.3	225	4	US-09-134-000C-4312	Sequence 4312, Ap
40	38	37.3	445	4	US-09-489-039A-13869	Sequence 13869, A
41	38	37.3	456	1	US-08-205-719-4	Sequence 4, Appl
42	38	37.3	456	3	US-08-431-517F-6	Sequence 6, Appl
43	38	37.3	482	3	US-08-431-517F-5	Sequence 5, Appl
44	38	37.3	482	3	US-08-431-517F-5	Sequence 5, Appl
45	38	37.3	505	4	US-09-627-216A-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-09-316-630-3
Sequence 3, Application US/09316630
Patent No. 6291637
GENERAL INFORMATION:
APPLICANT: Das, Saumitra
TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
TITLE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
FILE REFERENCE: 22000-20548-21
CURRENT APPLICATION NUMBER: US/09/316,630
CURRENT FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: 08/817,953
PRIOR FILING DATE: 1997-10-06
PRIOR APPLICATION NUMBER: 08/321,427
PRIOR FILING DATE: 1994-10-11
PRIOR APPLICATION NUMBER: 60/086,527
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 18
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3
Query Match 85.3%; Score 87; DB 3; Length 18;
Best local Similarity 100.0%; Pred. No. 8.7e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 4 EAKICHOIEYFGDF 18
DB 4 EAKICHOIEYFGDF 18
RESULT 2
US-09-316-630-4
Sequence 4, Application US/09316630
Patent No. 6291637
GENERAL INFORMATION:
APPLICANT: Das, Saumitra
TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
TITLE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
FILE REFERENCE: 22000-20548-21
CURRENT APPLICATION NUMBER: US/09/316,630
CURRENT FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: 08/817,953

Query Match 85.3%; Score 87; DB 7; Length 415;
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHQIERYFGDF 18
 |||||
 DB 14 EAKICHQIERYFGDF 28

RESULT 15
 ADE63993
 ID ADE63993 standard; protein; 415 AA.

AC ADE63993;

DT 29-JAN-2004 (first entry)

DE Rat Protein P38656, SEQ ID NO 9939.

KM Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

PN WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-033347P.

PA (GENO) GEN HOSPITAL CORP.

PI (FARB) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

DR WPI; 2003-268312/26.

DR GENBANK; P38656.

PT New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 415 AA;

Query Match 85.3%; Score 87; DB 7; Length 415;
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHQIERYFGDF 18
 |||||
 DB 14 EAKICHQIERYFGDF 28

Search completed: September 10, 2004, 17:51:46
 Job time : 57.4128 secs

CC the invention compete with LAP and inhibit the utilisation of various
 CC biochemical and physiological functions of LAP required for a productive
 CC life cycle. The methods and compositions are useful as antiviral therapy
 CC in pharmaceutical, veterinary or agricultural/horticultural applications.
 CC ABG72101-ABG72119 represent peptides useful as antiviral agents
 CC XX

SQL Sequence 18 AA;

Query Match 85.3%; Score 87; DB 6; Length 18;

Best Local Similarity 100.0%; Pred.No. 3.9e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EAKICHQIEYFGDF 18
 |||||
 Db 4 EAKICHQIEYFGDF 18

RESULT 7

AA01351

ID AAG01351 standard; protein; 92 AA.

AC AAG01351;

DT 06-OCT-2000 (first entry)

XX Human secreted protein, SEQ ID NO: 5432.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.

OS Homo sapiens.

PN EPI033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-00200610.

PR 26-FEB-1999; 99US-0122487P.

XX (GEST) GENSET.

PA Dumas Mline Edwards J, Duclert A, Giordano J;

PI WPI; 2000-500381/45.

DR N-PSDB; AAC01357.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

PS Claim 13; SEQ ID NO 5432; 71bp + Sequence Listing; English.

CC The present sequence is a polypeptide encoded by one of a large number of
 CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
 CC prepared from total human RNAs or polyA+ RNAs derived from 30 different
 CC tissues. EST sequences usually correspond mainly to the 3' untranslated
 CC region (UTR) of the mRNA because they are often obtained from oligo-dT
 CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
 CC sequences derived from the 5' ends of mRNAs and even in those cases where
 CC longer cDNA sequences have been obtained, the full 5' UTR is rarely
 CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can
 CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
 CC are also used in diagnostic, forensic, gene therapy and chromosome
 CC mapping procedures. They are used to obtain upstream regulatory sequences
 CC and to design expression and secretion vectors

SQL Sequence 92 AA;

Query Match 85.3%; Score 87; DB 3; Length 92;

Best Local Similarity 100.0%; Pred.No. 2.2e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EAKICHQIEYFGDF 18
 |||||
 Db 14 EAKICHQIEYFGDF 28

RESULT 8

AA03716

ID AAM03716 standard; protein; 408 AA.

AC AAM03716;

DT 25-MAR-2003 (revised)

DT 12-MAR-1997 (first entry)

XX Human autoantigen Ia(SS-B).

KW Autoimmune disease; Ia autoantigen; Sjogren's syndrome;
 KW systemic lupus erythematosus; diagnosis.

OS Homo sapiens.

PN US5541291-A.

PD 30-JUL-1996.

PF 27-MAY-1987; 87US-00054871.

PR 31-DEC-1984; 84US-00687908.

PA (UYDU-) UNIV DUKE.

PI Keene JD;

DR WPI; 1996-362015/36.

PT Auto-antigen U2-RNP, associated with severe polymyositis-scleroderma
 PT overlap syndrome - useful for diagnosis and treatment of autoimmune
 PT diseases.

PS Disclosure; Col 15-16; 21pp; English.

CC The human lupus antigen (Ia) is diagnostic for Sjogren's syndrome, as
 CC well as occurring in systemic lupus erythematosus patients. The Ia protein
 CC is clinically related to the Ro protein that is highly common among
 CC autoimmune patients. Ia and Ro antigens sometimes reside on the same
 CC cellular ribonucleoprotein particle; most Ia patients contain some Ro
 CC antibodies and vice versa. Ia cDNA has been isolated from a human liver
 CC library. (Updated on 25-MAR-2003 to correct PF field.)

SQL Sequence 408 AA;

Query Match 85.3%; Score 87; DB 2; Length 408;

Best Local Similarity 100.0%; Pred.No. 1.1e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EAKICHQIEYFGDF 18
 |||||
 Db 14 EAKICHQIEYFGDF 28

RESULT 9

ABP65252

ID ABP65252 standard; protein; 408 AA.

AC ABP65252;

DT 12-NOV-2002 (first entry)

XX Hypoxia-regulated protein #126.

KW Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;
 KW antiinflammatory; vulnery; gynecological; ophthalmological; vaccine;
 KW hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;

PD 02-DEC-1999 .
 PP 21-MAY-1999; 99WO-US011281.
 PR 22-MAY-1998; 98US-0086527P.
 RA (RBGC) UNIV CALIFORNIA.
 P1 Das S, Dasgupta A;
 DR WPI; 2000-062712/05.
 XX New yeast inhibitory peptide useful for inhibiting viral protein
 FT translation and replication.
 XX
 PS Claim 5; Page 57; 81pp; English.
 XX
 CC This sequence is the La autoantigen binding domain (LAP). LAP is a
 CC cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 CC translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors from
 CC binding at the site. The peptide can be used alone or in combination with
 CC an inhibitor RNA (IRNA see RA245200). The LAP peptide is useful as an
 CC antiviral agent, which works through the inhibition of mRNA translation,
 CC especially viral mRNA. Examples of viruses which can be inhibited are
 CC rhinovirus, flavivirus, coronavirus, hepatitis A B or C viruses,
 CC hepadornavirus, adenovirus, and parainfluenza virus, poliovirus, rhinovirus,
 CC coxsackie virus, encephalomyocarditis virus, foot-and-mouth disease
 CC virus, echo virus, hepatitis C virus, infectious bronchitis virus, duck
 CC and human hepatitis B virus, and vesicular stomatitis virus. The peptide
 CC also inhibits replication of the above viruses. The LAP peptide
 CC selectively inhibits viral protein translation, and is therefore not
 CC toxic to the host cell
 XX
 SQ Sequence 18 AA;

```

Query Match      85.3%; Score 87; DB 3; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY          4 EAKICHQIEYYRGDF 18
            |||||
Db           4 EAKICHQIEYYRGDF 18

RESULT 5
ABG72101
ID ABG72101 standard; peptide; 18 AA.
XX AC          ABG72101;
XX DT          28-JAN-2003 (first entry)
XX DE          Viral replication inhibiting peptide, LAP.
XX KM          Viral replication inhibitor; IRMS initiated translation; LAP;
KM internal ribosome entry site initiated translation; La antigen protein;
KM viral life cycle; antiviral therapy; pharmaceutical; veterinary;
KM agricultural; horticultural; virucide.
OS Unidentified.
XX PN          WO200283858-A2.
XX PD          24-OCT-2002.
XX PF          12-APR-2002; 2002WO-US011589.
XX PR          16-APR-2001; 2001US-00836073.
XX (UYCA-) UNITV CALIFORNIA OFFICE PRESIDENT.
```

XX	Dasgupta A,	Dae S,	Baidya N,
PI	WPI;	2003-058634/05.	
DR			
XX	New compound containing acidic and aromatic amino acids, useful as		
PT	antiviral therapy in pharmaceutical, veterinary or		
PT	agricultural/horticultural applications.		
XX			
PS	Claim 1; Page 15; 19pp; English.		
XX			
CC	The present invention relates to peptides and methods of inhibiting the		
CC	replication of viruses that utilise internal ribosome entry site (IRES)		
CC	initiated translation, and/or inhibiting viruses that utilise the la		
CC	antigen protein (LAP) in any phase of their life cycle. The peptides of		
CC	the invention compete with LAP and inhibit the utilisation of various		
CC	biochemical and physiological functions of LAP required for a productive		
CC	life cycle. The methods and compositions are useful as antiviral therapy		
CC	in pharmaceutical, veterinary or agricultural/horticultural applications.		
CC	ABG72101-ABG72119 represent peptides useful as antiviral agents		
SQ			
	Sequence 18 AA;		
OY	Query Match	85.3%; Score 87; DB 6; Length 18;	
	Best Local Similarity	100.0%; Pred. No. 3.9e-07;	
DB	Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
	4 EAKICHOIEVYFGDF 18		
	4 EAKICHOIEVYFGDF 18		

RESULT 6
 ABG72114
 ID ABG72114 standard; peptide; 18 AA.
 XX
 AC ABG72114;
 XX
 DT 28-JAN-2003 (first entry)
 XX
 DE Viral replication inhibiting peptide; BOVINE.
 XX
 KW Viral replication inhibitor; IRBS initiated translation; LAP;
 KW internal ribosome entry site initiated translation; Va antigen protein;
 KW viral life cycle; antiviral therapy; pharmaceutical; veterinary;
 KW agricultural; horticultural; virucide; bovine.
 XX
 OS *Sovinae*.
 XX
 PN WO200283858-A2.
 XX
 PD 24-OCT-2002.
 XX
 PF 12-APR-2002; 2002WO-US011589.
 XX
 PR 16-APR-2001; 2001US-00836073.
 XX
 PA (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.
 XX
 PI Dasgupta A, Das S, Balda N;
 XX
 DR WPI; 2003-058634/05.
 XX
 PT New compound containing acidic and aromatic amino acids, useful as
 PT antiviral therapy in pharmaceutical, veterinary or
 PT agricultural/horticultural applications.
 XX
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC The present invention relates to peptides and methods of inhibiting the
 CC replication of viruses that utilise internal ribosome entry site (IRES)
 CC initiated translation, and/or inhibiting viruses that utilise the La
 CC antigen protein (LAP) in any phase of their life cycle. The peptides of

Query Match 100.0%; Score 102; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 1.4e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 000EAKICHQIEYFGDF 18
 |||||
 Db 1 000EAKICHQIEYFGDF 18

RESULT 2
 ABG72105
 ID ABG72105 standard; peptide; 18 AA.
 AC ABG72105;
 XX
 XX
 DT 28-JAN-2003 (first entry)
 XX
 DE Viral replication inhibiting peptide, 762.
 XX
 XX Viral replication inhibitor; IRES initiated translation; LAP;
 KM internal ribosome entry site initiated translation; La antigen protein;
 KM viral life cycle; antiviral therapy; pharmaceutical; veterinary;
 KM agricultural; horticultural; virucide.
 XX
 OS Unidentified.
 XX
 PN WO200283858-A2.
 XX
 PD 24-OCT-2002.
 XX
 PF 12-APR-2002; 2002WO-US011589.
 XX
 PR 16-APR-2001; 2001US-00836073.
 XX
 PA (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.
 XX
 PI Dasgupta A, Das S, Baidya N;
 XX
 DR MPI; 2003-058634/05.
 XX
 PT New compound containing acidic and aromatic amino acids, useful as
 PT antiviral therapy in pharmaceutical, veterinary or
 PT agricultural/horticultural applications.
 PT
 XX
 PS Example 3; Page 14; 19pp; English.
 XX
 CC The present invention relates to peptides and methods of inhibiting the
 CC replication of viruses that utilise internal ribosome entry site (IRES)
 CC initiated translation, and/or inhibiting viruses that utilise the La
 CC antigen protein (LAP) in any phase of their life cycle. The peptides of
 CC the invention compete with LAP and inhibit the utilisation of various
 CC biochemical and physiological functions of LAP required for a productive
 CC life cycle. The methods and compositions are useful as antiviral therapy
 CC in pharmaceutical, veterinary or agricultural/horticultural applications.
 CC ABG72101-ABG72119 represent peptides useful as antiviral agents
 CC
 XX
 SQ Sequence 18 AA;

Query Match 88.2%; Score 90; DB 6; Length 18;
 Best Local Similarity 88.9%; Pred. No. 1.3e-07;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 000EAKICHQIEYFGDF 18
 |||||
 Db 1 000EAKICHQIEYFGDF 18

RESULT 3
 ABG72113
 ID ABG72113 standard; peptide; 17 AA.
 AC ABG72113;
 XX

XX
 DT 28-JAN-2003 (first entry)
 XX
 DE Viral replication inhibiting peptide, MOUSE.
 XX
 XX Viral replication inhibitor; IRES initiated translation; LAP;
 KM internal ribosome entry site initiated translation; La antigen protein;
 KM viral life cycle; antiviral therapy; pharmaceutical; veterinary;
 KM agricultural; horticultural; virucide; mouse.
 XX
 OS Mus sp.
 XX
 PN WO200283858-A2.
 XX
 PD 24-OCT-2002.
 XX
 PF 12-APR-2002; 2002WO-US011589.
 XX
 PR 16-APR-2001; 2001US-00836073.
 XX
 PA (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.
 XX
 PI Dasgupta A, Das S, Baidya N;
 XX
 DR MPI; 2003-058634/05.
 XX
 PT New compound containing acidic and aromatic amino acids, useful as
 PT antiviral therapy in pharmaceutical, veterinary or
 PT agricultural/horticultural applications.
 PT
 XX
 PS Claim 1; Page 15; 19pp; English.
 XX
 CC The present invention relates to peptides and methods of inhibiting the
 CC replication of viruses that utilise internal ribosome entry site (IRES)
 CC initiated translation, and/or inhibiting viruses that utilise the La
 CC antigen protein (LAP) in any phase of their life cycle. The peptides of
 CC the invention compete with LAP and inhibit the utilisation of various
 CC biochemical and physiological functions of LAP required for a productive
 CC life cycle. The methods and compositions are useful as antiviral therapy
 CC in pharmaceutical, veterinary or agricultural/horticultural applications.
 CC ABG72101-ABG72119 represent peptides useful as antiviral agents
 CC
 XX
 SQ Sequence 17 AA;

Query Match 85.3%; Score 87; DB 6; Length 17;
 Best Local Similarity 100.0%; Pred. No. 3.6e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHQIEYFGDF 18
 |||||
 Db 3 EAKICHQIEYFGDF 17

RESULT 4
 AA552200
 ID AA552200 standard; peptide; 18 AA.
 AC AA552200;
 XX
 XX
 DT 14-MAR-2000 (first entry)
 XX
 DE Human La autoantigen peptide (LAP).
 XX
 XX La autoantigen; LAP; internal ribosome entry site; IRES; translation;
 KM viral replication; RNA; antiviral agent; picornavirus; flavivirus;
 KM coronavirus; hepatitis virus; rhabdovirus; adenovirus; coxsackie virus;
 KM parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
 KM foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
 KM vesicular stomatitis virus.
 XX
 OS Homo sapiens.
 XX
 PN WO9961613-A2.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 16:58:20 ; Search time 56.3129 Seconds
(without alignments)
90.314 Million cell updates/sec

Title: US-09-836-073-4

Perfect score: 102
Sequence: 1 QOOEAKICHIQIHYEGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1980s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	102	100.0	18	6	ABG72104 Viral rep
2	90	88.2	18	6	ABG72105 Viral rep
3	87	85.3	17	6	ABG72113 Viral rep
4	87	85.3	18	3	AAV52200 Human Ia
5	87	85.3	18	6	ABG72101 Viral rep
6	87	85.3	18	6	ABG72114 Viral rep
7	87	85.3	92	3	AAAG01351 Human sec
8	87	85.3	408	2	AAW03716 Human aut
9	87	85.3	408	5	ABP65252 Hypoxia-x
10	87	85.3	408	5	ABP65252 Human pro
11	87	85.3	408	7	ADDD46272 Human pro
12	87	85.3	408	7	ADDE63995 Human pro
13	87	85.3	408	7	ADDE62859 Human pro
14	87	85.3	415	7	ADDE63989 Rat Prote
15	87	85.3	415	7	ADDE63993 Rat Prote
16	87	85.3	439	4	ABG08417 Novel hum
17	87	85.3	460	3	AAAS8987 Breast an
18	87	85.3	460	5	ABP41511 Human ova
19	81	79.4	18	6	ABG72109 Viral rep
20	79	77.5	18	6	ABG72111 Viral rep
21	79	77.5	18	6	ABG72112 Viral rep
22	78	76.5	18	6	ABG72110 Viral rep
23	76.5	75.0	19	6	ABG72116 Viral rep
24	75	73.5	18	6	ABG72102 Viral rep
25	75	73.5	18	6	ABG72103 Viral rep

26	73	71.6	21	2	AA43394 Ia/SSB ep
27	72	70.6	18	6	ABG72115 Viral rep
28	72	70.6	18	6	ABG72107 Viral rep
29	71	69.6	18	6	ABG72108 Viral rep
30	62	60.8	16	6	ABG72119 Viral rep
31	57	55.9	390	4	ABBE5316 Drosophi1
32	56	54.9	18	6	ABG72106 Viral rep
33	52	51.0	381	6	ABJ19009 Patchogen
34	52	51.0	391	6	ABM72091 Staphyloc
35	48	47.1	913	3	AA647714 Arabidops
36	48	47.1	923	3	AA647712 Arabidops
37	48	47.1	993	3	AA647712 Arabidops
38	47	46.1	411	6	ABP70526 Histone d
39	46	45.1	33	7	ABO23427 Amino aci
40	46	45.1	377	6	ABU43510 Protein e
41	45	44.1	1900	4	ABBS9107 Drosophi1
42	44	43.1	18	6	ABG72117 Viral rep
43	44	43.1	135	4	ABG27058 Novel hum
44	44	43.1	224	5	ABBB9645 Human pol
45	44	43.1	224	6	ADA55336 Human pro

ALIGNMENTS

RESULT 1
ID ABG72104 standard; peptide; 18 AA.
AC ABG72104;
DT 28-JAN-2003 (first entry)
DE Viral replication inhibiting peptide, 761.
XX Viral replication inhibitor; IRIS initiated translation; LAP;
XX internal ribosome entry site initiated translation; La antigen protein;
XX viral life cycle; antiviral therapy; pharmaceutical; veterinary;
XX agricultural; horticultural; virucide.
OS Unidentified.
OS
XX
XX WO200283858-A2.
XX
XX 24-OCT-2002.
XX
XX 12-APR-2002; 2002WO-US011589.
XX
XX 16-APR-2001; 2001US-00836073.
XX
XX (UYCA-) UNIV CALIFORNIA OFFICE PRESIDENT.
XX
XX Dasgupta A, Das S, Baidya N;
XX
XX WPI; 2003-058634/05.
XX
XX
XX New compound containing acidic and aromatic amino acids, useful as
XX antiviral therapy in pharmaceutical, veterinary or
XX agricultural/horticultural applications.
XX
XX Claim 10; Page 16; 19pp; English.
XX
XX The present invention relates to peptides and methods of inhibiting the
XX replication of viruses that utilize internal ribosome entry site (IRES)
XX initiated translation, and/or inhibiting viruses that utilize the La
XX antigen protein (LAP) in any phase of their life cycle. The peptides of
XX the invention compete with LAP and inhibit the translation of various
XX biochemical and physiological functions of LAP required for a productive
XX life cycle. The methods and compositions are useful as antiviral therapy
XX in pharmaceutical, veterinary or agricultural/horticultural applications.
XX ABG72101-ABG72119 represent peptides useful as antiviral agents
XX
XX Sequence 18 AA;

This Page Blank (uspio)

Search completed: September 10, 2004, 18:11:55
Job time : 41.424 secs

```
FILE REFERENCE: 220002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 18
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match      80.6%; Score 79; DB 9; Length 18;
Best Local Similarity 77.8%; Pred. No. 8.3e-06;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
Db 1 AALEAKICHQIEYFGDF 18

RESULT 12
US-09-836-073-10
Sequence 10, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Das, S.
APPLICANT: Balidya, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 220002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 18
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-836-073-10

Query Match      79.6%; Score 78; DB 9; Length 18;
Best Local Similarity 77.8%; Pred. No. 1.2e-05;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
Db 1 AALEAKICHQIEYFGDF 18

RESULT 13
US-09-836-073-16
Sequence 16, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Das, S.
APPLICANT: Balidya, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 220002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 19
TYPE: PRT
ORGANISM: Rat
```

```
US-09-836-073-16

Query Match      78.1%; Score 76.5; DB 9; Length 19;
Best Local Similarity 78.9%; Pred. No. 2.2e-05;
Matches 15; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 AALQAKICHQIQYFGQF 18
Db 1 AALEAKICHQIEYFGDF 19

RESULT 14
US-09-836-073-2
Sequence 2, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Das, S.
APPLICANT: Balidya, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 220002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 18
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match      76.5%; Score 75; DB 9; Length 18;
Best Local Similarity 72.2%; Pred. No. 3.7e-05;
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
Db 1 AALEAKICHQIEYFGDF 18

RESULT 15
US-09-836-073-4
Sequence 4, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Das, S.
APPLICANT: Balidya, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 220002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 18
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-836-073-4

Query Match      76.5%; Score 75; DB 9; Length 18;
Best Local Similarity 80.0%; Pred. No. 3.7e-05;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 QAKICHQIQYFGQF 18
Db 4 EAKICHQIEYFGDF 18
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-806-695

Query Match      88.8%; Score 87; DB 14; Length 460;
Best Local Similarity 83.3%; Pred. No. 1.1e-05;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 AALQAKICHQIQYFFGQF 18
DB      63 AALEAKICHQIEYFGDF 80

RESULT 7
US-10-264-049-2643
; Sequence 2643, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2643
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-2643

Query Match      88.8%; Score 87; DB 15; Length 460;
Best Local Similarity 83.3%; Pred. No. 1.1e-05;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 AALQAKICHQIQYFFGQF 18
DB      63 AALEAKICHQIEYFGDF 80

RESULT 8
US-09-836-073-13
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Mouse
US-09-836-073-13

Query Match      84.7%; Score 83; DB 9; Length 17;
Best Local Similarity 82.4%; Pred. No. 1.8e-06;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      2 ALOAKICHQIQYFFGQF 18
DB      1 ALEAKICHQIEYFGDF 17
```

```
RESULT 9
US-09-836-073-9
; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-9

Query Match      82.7%; Score 81; DB 9; Length 18;
Best Local Similarity 87.5%; Pred. No. 3.9e-06;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 AALQAKICHQIQYFFG 16
DB      1 AALEAKICHQIEYFG 16
```

```
RESULT 10
US-09-836-073-11
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-11

Query Match      80.6%; Score 79; DB 9; Length 18;
Best Local Similarity 77.8%; Pred. No. 8.3e-06;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 AALQAKICHQIQYFFGQF 18
DB      1 AALEAKICHQIEYFGDF 18

RESULT 11
US-09-836-073-12
; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
```


APPLICANT: Baliga, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 220002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 18
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-836-073-1

Query Match 88.8%; Score 87; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 4.3e-07;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
DB 1 AALEAKICHQIEYFGDF 18

RESULT 3
US-09-836-073-14
Sequence 14, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Das, S.
APPLICANT: Baliga, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 220002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 18
TYPE: PRT
ORGANISM: Bovine
US-09-836-073-14

Query Match 88.8%; Score 87; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 4.3e-07;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
DB 1 AALEAKICHQIEYFGDF 18

RESULT 4
US-10-170-385-477
Sequence 477, Application US/10170385
Publication No. US2003020372A1
GENERAL INFORMATION:
APPLICANT: Ward, Neil Raymond
APPLICANT: Mundy, Christopher Robert
APPLICANT: Kan, On
APPLICANT: Harris, Robert Alan
APPLICANT: White, Jonathan
APPLICANT: Binley, Katie Mary
APPLICANT: Rayner, William Nigel
APPLICANT: Naylor, Stuart
APPLICANT: Kingsman, Susan Mary
APPLICANT: Krige, David
TITLE OF INVENTION: ANALYSIS METHOD
FILE REFERENCE: 532682000100
CURRENT APPLICATION NUMBER: US/10/170,385

CURRENT FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: PCT/GB02/01662
PRIOR FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: PCT/GB01/05458
PRIOR FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 549
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 477
LENGTH: 408
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-170-385-477

Query Match 88.8%; Score 87; DB 12; Length 408;
Best Local Similarity 83.3%; Pred. No. 9.7e-06;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
DB 11 AALEAKICHQIEYFGDF 28

RESULT 5
US-09-925-298-695
Sequence 695, Application US/09925298
Publication No. US20020039764A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA103
CURRENT APPLICATION NUMBER: US/09/925,298
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05881
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 846
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 695
LENGTH: 460
TYPE: PRT
ORGANISM: Homo sapiens
US-09-925-298-695

Query Match 88.8%; Score 87; DB 12; Length 460;
Best Local Similarity 83.3%; Pred. No. 1.1e-05;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
DB 63 AALEAKICHQIEYFGDF 80

RESULT 6
US-10-102-806-695
Sequence 695, Application US/10102806
Publication No. US20030054421A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA103P1
CURRENT APPLICATION NUMBER: US/10/102,806
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/925,298
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05881
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 846
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 695
LENGTH: 460

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:52:06 ; Search time 40.324 Seconds
(without alignments)
143.151 Million cell updates/sec

Title: US-09-836-073-3.

Perfect score: 98
Sequence: 1 AALQAKICHQIQYFGQF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCF_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PCFUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09C_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	100.0	18	9	US-09-836-073-3
2	87	88.8	18	9	US-09-836-073-1
3	87	88.8	18	9	US-09-836-073-14
4	87	88.8	408	12	US-10-170-385-477
5	87	88.8	460	12	US-09-925-298-695
6	87	88.8	460	14	US-10-102-806-695
7	87	88.8	460	15	US-10-264-049-2643
8	83	84.7	17	9	US-09-836-073-13
9	81	82.7	18	9	US-09-836-073-9
10	79	80.6	18	9	US-09-836-073-11
11	79	80.6	18	9	US-09-836-073-12
12	78	79.6	18	9	US-09-836-073-10
13	76.5	78.1	19	9	US-09-836-073-16
14	75	76.5	18	9	US-09-836-073-2
15	75	76.5	18	9	US-09-836-073-4

16	72	73.5	18	9	US-09-836-073-7	Sequence 7, Appli
17	71	72.4	18	9	US-09-836-073-8	Sequence 8, Appli
18	65	66.3	18	9	US-09-836-073-15	Sequence 15, Appli
19	64	65.3	21	15	US-10-376-121A-20	Sequence 20, Appli
20	63	64.3	18	9	US-09-836-073-5	Sequence 5, Appli
21	59	60.2	38	12	US-10-325-810-214	Sequence 214, App
22	56	57.1	18	9	US-09-836-073-6	Sequence 6, Appli
23	52	53.1	420	16	US-10-437-963-128072	Sequence 128072,
24	50	51.0	38	9	US-09-843-676-25	Sequence 25, Appli
25	50	51.0	38	9	US-09-766-253-25	Sequence 25, Appli
26	50	51.0	38	10	US-09-438-486-25	Sequence 25, Appli
27	50	51.0	38	12	US-10-325-810-215	Sequence 215, App
28	50	51.0	38	14	US-10-053-758-25	Sequence 25, Appli
29	50	51.0	38	14	US-10-054-295-25	Sequence 25, Appli
30	50	51.0	38	14	US-10-054-611-25	Sequence 25, Appli
31	49.5	50.5	37	9	US-09-843-676-24	Sequence 24, Appli
32	49.5	50.5	37	9	US-09-766-253-24	Sequence 24, Appli
33	49.5	50.5	37	10	US-09-438-486-24	Sequence 24, Appli
34	49.5	50.5	37	14	US-10-053-758-24	Sequence 24, Appli
35	49.5	50.5	37	14	US-10-054-295-24	Sequence 24, Appli
36	49.5	50.5	37	14	US-10-054-611-24	Sequence 24, Appli
37	49	50.0	862	16	US-10-437-963-110390	Sequence 130390,
38	47	48.0	609	10	US-09-805-455-4	Sequence 4, Appli
39	47	48.0	609	16	US-10-678-140-4	Sequence 4, Appli
40	47	48.0	615	10	US-09-805-455-2	Sequence 2, Appli
41	47	48.0	615	16	US-10-678-140-2	Sequence 2, Appli
42	47	48.0	664	10	US-09-805-455-5	Sequence 5, Appli
43	47	48.0	664	16	US-10-678-140-5	Sequence 5, Appli
44	46	46.9	377	12	US-10-282-122A-71434	Sequence 71434, A
45	46	46.9	608	10	US-09-805-455-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-836-073-3
; Sequence 3, Application US/09836073
; Patient No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ. ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; TYPE: PRF
; LENGTH: 18
; ORGANISM: Homo Sapiens
US-09-836-073-3

Query Match 100.0%; Score 98; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 7.2e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AALQAKICHQIQYFGQF 18
Db 1 AALQAKICHQIQYFGQF 18
RESULT 2
US-09-836-073-1
; Sequence 1, Application US/09836073
; Patient No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-24

Query Match 50.5%; Score 49.5; DB 3; Length 37;
Best Local Similarity 75.0%; Pred. No. 0.2;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICHQIQYFFGQF 18
|||:||||
1 ICHQ-EYFGDF 11

Db

RESULT 15
US-09-430-323-24
Sequence 24, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin
Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-430-323-24

Query Match 50.5%; Score 49.5; DB 4; Length 37;
Best Local Similarity 75.0%; Pred. No. 0.2;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICHQIQYFFGQF 18
|||:||||
1 ICHQ-EYFGDF 11

Db

Search completed: September 10, 2004, 18:05:07
Job time: 15.8883 secs

FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0300
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 215:
US-09-721-456-215

Query Match 51.0%; Score 50; DB 4; Length 38;
Best Local Similarity 80.0%; Pred. No. 0.17;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIOXYFG 16
Db 1 ICEQIEYFG 10

RESULT 13
US-08-851-843A-24
Sequence 24, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-24

Query Match 50.5%; Score 49.5; DB 3; Length 37;
Best Local Similarity 75.0%; Pred. No. 0.2;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICHQIOXYFGQF 18
Db 1 ICHQ-EYFGDF 11

RESULT 14
US-08-854-050-24
Sequence 24, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6261836el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536

FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-430-323-25

Query Match 51.0%; Score 50; DB 4; Length 38;
Best Local Similarity 80.0%; Pred. No. 0.17;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIQYFYG 16
|||:||||
DB 1 ICEQIEYFYG 10

RESULT 11
US-09-402-181B-215
Sequence 215, Application US/09402181B
Patent No. 6610839
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 633
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,181B
FILING DATE: 29-Sep-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/117885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ausenhue, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 215:
US-09-402-181B-215

Query Match 51.0%; Score 50; DB 4; Length 38;
Best Local Similarity 80.0%; Pred. No. 0.17;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIQYFYG 16
|||:||||
DB 1 ICEQIEYFYG 10

RESULT 12
US-09-721-456-215
Sequence 215, Application US/09721456
Patent No. 6617110
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/721,456
FILING DATE: 22-NOV-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017

FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17865
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-215

Query Match 51.0%; Score 50; DB 3; Length 38;
Best Local Similarity 80.0%; Pred. No. 0.17;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 7 ICHQIOYFYFG 16
|||:||||
Db 1 ICHQIEYFYFG 10

RESULT 9
US-08-854-050-25
Sequence 25, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6261836el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-25

Query Match 51.0%; Score 50; DB 3; Length 38;
Best Local Similarity 80.0%; Pred. No. 0.17;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 7 ICHQIOYFYFG 16
|||:||||
Db 1 ICHQIEYFYFG 10

RESULT 10
US-09-430-323-25
Sequence 25, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843

FILED DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 214:
US-09-721-456-214

Query Match 60.2%; Score 59; DB 4; Length 38;
Best Local Similarity 75.0%; Pred. No. 0.0059;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHQIOYFGQF 18
DB 1 ICHQXEYFGDF 12

RESULT 7
US-08-851-843A-25
Sequence 25, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-25

Query Match 51.0%; Score 50; DB 3; Length 38;
Best Local Similarity 80.0%; Pred. No. 0.17;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIOYFG 16
DB 1 ICHQXEYFG 10

RESULT 8
US-08-974-549A-215
Sequence 215, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050

```

; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 214:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
US-08-974-549A-214

Query Match          60.2%; Score 59; DB 3; Length 38;
Best Local Similarity 75.0%; Pred. No. 0.0059;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHQIQYFGQF 18
Db 1 ICHQXEYFGDF 12

RESULT 5
US-09-402-181B-214
; Sequence 214, Application US/09402181B
; Patent No. 6610839
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
;             Lingner, Joachim
;             Nakamura, Toru
;             Chapman, Karen B.
;             Morin, Gregg B.
;             Harley, Calvin B.
;             Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 633
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/402,181B
; FILING DATE: 29-Sep-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: US 08/915,503
```

```

; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ausubius, Scott L.
; REGISTRATION NUMBER: 42,271
; REFERENCE/DOCKET NUMBER: 015389-002620US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 214:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 214:
;
US-09-402-181B-214

Query Match          60.2%; Score 59; DB 4; Length 38;
Best Local Similarity 75.0%; Pred. No. 0.0059;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHQIQYFGQF 18
Db 1 ICHQXEYFGDF 12

RESULT 6
US-09-721-456-214
; Sequence 214, Application US/09721456
; Patent No. 6617110
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
;             Lingner, Joachim
;             Nakamura, Toru
;             Chapman, Karen B.
;             Morin, Gregg B.
;             Harley, Calvin B.
;             Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/721,456
; FILING DATE: 22-NOV-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/911,312
```


PRIOR FILING DATE: 1997-10-06
PRIOR APPLICATION NUMBER: 08/321,427
PRIOR FILING DATE: 1994-10-11
PRIOR APPLICATION NUMBER: 60/086,527
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO: 4
LENGTH: 18
TYPE: PR1
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: B-LAP
US-09-316-630-4

Query Match 88.8%; Score 87; DB 3; Length 18;
Best Local Similarity 83.3%; Pred. No. 7,5e-08;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AALQAKICHQIQYFGQF 18
Db 1 AALEAKICHQIEYFGDF 18

RESULT 3
US-08-475-955-20
Sequence 20, Application US/08475955
Patent No. 6641813
GENERAL INFORMATION:
APPLICANT: Harley, John
TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS OF
NUMBER OF SEQUENCES: 218
CORRESPONDENCE ADDRESS:
ADDRESSER: Patricia L. Pabst
STREET: 2800 One Atlantic Center, 1250 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,955
FILING DATE: June 7, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/867,819
FILING DATE: April 13, 1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/648,205
FILING DATE: January 31, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/472,947
FILING DATE: January 31, 1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patricia L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMR114CIP(2) DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-873-8795
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Binding-site
LOCATION: 8..15
US-08-475-955-20

Query Match 65.3%; Score 64; DB 4; Length 21;
Best Local Similarity 83.3%; Pred. No. 0.00049;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ICHQIQYFGQF 18
Db 1 ICHQIEYFGDF 12

RESULT 4
US-08-974-549A-214
Sequence 214, Application US/08974549A
Patent No. 616178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:43:01 ; Search time 15.8883 Seconds
(without alignments)
58.488 Million cell updates/sec

Title: US-09-836-073-3

Sequence: 1 AALQAKICHQIOYFGQF 18

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	88.8	18	3	US-09-316-630-3
2	87	88.8	18	3	US-09-316-630-4
3	64	65.3	21	4	US-08-475-955-20
4	59	60.2	38	3	US-08-974-549A-214
5	59	60.2	38	4	US-09-402-181B-214
6	59	60.2	38	4	US-09-721-456-214
7	50	51.0	38	3	US-08-851-843A-25
8	50	51.0	38	3	US-08-974-549A-215
9	50	51.0	38	3	US-08-854-050-25
10	50	51.0	38	4	US-09-430-323-25
11	50	51.0	38	4	US-09-402-181B-215
12	50	51.0	38	4	US-09-721-456-215
13	49.5	50.5	37	3	US-08-851-843A-24
14	49.5	50.5	37	3	US-08-854-050-24
15	49.5	50.5	37	4	US-09-430-323-24
16	42	42.9	1220	2	US-08-680-326-38
17	42	42.9	2644	4	US-09-029-047C-2
18	42	42.9	292	4	US-09-134-001C-3515
19	40	40.8	111	4	US-09-257-179-61
20	40	40.8	198	4	US-09-328-352-6436
21	40	40.8	554	4	US-09-540-236-3778
22	40	40.8	616	3	US-09-298-367B-11
23	39	39.8	39	3	US-08-851-843A-26
24	39	39.8	39	3	US-08-974-549A-216
25	39	39.8	39	3	US-08-854-050-26
26	39	39.8	39	4	US-09-430-323-26
27	39	39.8	39	4	US-09-402-181B-216

28	39	39.8	39	4	US-09-721-456-216	Sequence 216, App
29	39	39.8	502	4	US-09-328-352-5891	Sequence 5891, Ap
30	39	39.8	1336	4	US-09-883-134-4	Sequence 1, Appl
31	39	39.8	2954	4	US-09-150-867-1	Sequence 23, Appl
32	38.5	39.3	740	4	US-09-323-872A-23	Sequence 15, Appl
33	38.5	39.3	740	4	US-09-072-433-15	Sequence 28, Appl
34	38.5	39.3	864	4	US-09-323-872A-28	Sequence 16, Appl
35	38.5	39.3	864	4	US-09-072-433-16	Sequence 6934, Ap
36	38	38.8	185	4	US-09-107-532A-6934	Sequence 6229, Ap
37	38	38.8	289	4	US-09-328-352-6229	Sequence 25, Appl
38	38	38.8	305	1	US-08-420-235B-25	Sequence 25, Appl
39	38	38.8	305	3	US-08-793-624-25	Sequence 25, Appl
40	38	38.8	305	5	PCT-US95-10194-25	Sequence 6745, Ap
41	38	38.8	463	4	US-09-543-681A-6745	Sequence 30245, A
42	38	38.8	576	4	US-09-252-991A-30245	Sequence 2, Appl
43	38	38.8	984	4	US-09-287-354-2	Sequence 3, Appl
44	38	38.8	1189	4	US-09-287-354-3	Sequence 4, Appl
45	38	38.8	1189	4	US-09-287-354-4	

ALIGNMENTS

RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Dae, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: IAP
US-09-316-630-3
Query Match 88.8%; Score 87; DB 3; Length 18;
Best Local Similarity 83.3%; Pred. No. 7.5e-09;
Matches 15; Conservative 2; Mismatches 1; Indels 0;
Gaps 0;
Db 1 AALQAKICHQIOYFGQF 18
1 AALEAKICHQIERYFGDF 18
RESULT 2
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Dae, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953

This Page Blank (uspto)

Mon Sep 13 09:36:20 2004

us-09-836-073-2.rapb

Page 5

Search completed: September 10, 2004, 18:11:54
Job time : 40.424 secs

```

FILE REFERENCE: 220002054822
;
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match      82.3%; Score 79; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 1e-05;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 AALEAQCQIQIEYFGDF 18
      |||||:|||||
Db      1 AALEAKICHOIEYFGDF 18

RESULT 12
US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-10

Query Match      81.2%; Score 78; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 1.5e-05;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 AALEAQCQIQIEYFGDF 18
      |||||:|||||
Db      1 AALEAKICHOIEYFGDF 18

RESULT 13
US-09-836-073-16
; Sequence 16, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Rat

```

```

US-09-836-073-16

Query Match      79.7%; Score 76.5; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 2.7e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY      1 AALEAQCQIQI-EYFGDF 18
      |||||:|||||
Db      1 AALEAKICHOIEYFGDF 19

RESULT 14
US-09-836-073-3
; Sequence 3, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-3

Query Match      78.1%; Score 75; DB 9; Length 18;
Best Local Similarity 72.2%; Pred. No. 4.4e-05;
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 AALEAQCQIQIEYFGDF 18
      |||||:|||||
Db      1 AALQAKICHOIQYFGDF 18

RESULT 15
US-09-836-073-4
; Sequence 4, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-4

Query Match      78.1%; Score 75; DB 9; Length 18;
Best Local Similarity 86.7%; Pred. No. 4.4e-05;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 EAQICQIQIEYFGDF 18
      |||||:|||||
Db      4 EAKICHOIEYFGDF 18

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-806-695

Query Match          90.6%; Score 87; DB 14; Length 460;
Best Local Similarity 88.9%; Pred. No. 1.6e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQCQIQIEYFGDF 18
    |||||:|||||
Db 63 AALEAKICHQIEYFGDF 80

RESULT 7
US-10-264-049-2643
; Sequence 2643, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birtse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2643
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-2643

Query Match          90.6%; Score 87; DB 15; Length 460;
Best Local Similarity 88.9%; Pred. No. 1.6e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQCQIQIEYFGDF 18
    |||||:|||||
Db 63 AALEAKICHQIEYFGDF 80

RESULT 8
US-09-836-073-13
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baladya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Mouse
US-09-836-073-13

Query Match          86.5%; Score 83; DB 9; Length 17;
Best Local Similarity 88.2%; Pred. No. 2.2e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALEAQCQIQIEYFGDF 18
    |||||:|||||
Db 1 ALEAKICHQIEYFGDF 17
```

```

RESULT 9
US-09-836-073-9
; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baladya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-073-9

Query Match          84.4%; Score 81; DB 9; Length 18;
Best Local Similarity 88.2%; Pred. No. 4.8e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQCQIQIEYFGDF 17
    |||||:|||||
Db 1 AALEAKICHQIEYFGDF 17

RESULT 10
US-09-836-073-11
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baladya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-073-11

Query Match          82.3%; Score 79; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 1e-05;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALEAQCQIQIEYFGDF 18
    |||||:|||||
Db 1 AALEAKICHQIEYFGDF 18

RESULT 11
US-09-836-073-12
; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baladya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
```

```

; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1
; LENGTH: 18
; TYPE: PRN
; ORGANISM: Homo Sapiens
US-09-836-073-1

Query Match          90.6%; Score 87; DB 9; Length 18;
Best Local Similarity 88.9%; Pred. No. 5.3e-07;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 AALEAQCQCIIEYFGDF 18
        |||||:|||||
DB      1 AALEAKICHOIEYFGDF 18

RESULT 3
US-09-836-073-14
; Sequence 14, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 14
; LENGTH: 18
; TYPE: PRN
; ORGANISM: Bovine
US-09-836-073-14

Query Match          90.6%; Score 87; DB 9; Length 18;
Best Local Similarity 88.9%; Pred. No. 5.3e-07;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 AALEAQCQCIIEYFGDF 18
        |||||:|||||
DB      1 AALEAKICHOIEYFGDF 18

RESULT 4
US-10-170-385-477
; Sequence 477, Application US/10170385
; Publication No. US20030203372A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingeman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 532682000100
; CURRENT APPLICATION NUMBER: US/10/170,385
```

```

; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 477
; LENGTH: 408
; TYPE: PRN
; ORGANISM: Homo Sapiens
US-10-170-385-477

Query Match          90.6%; Score 87; DB 12; Length 408;
Best Local Similarity 88.9%; Pred. No. 1.4e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 AALEAQCQCIIEYFGDF 18
        |||||:|||||
DB      11 AALEAKICHOIEYFGDF 28

RESULT 5
US-09-925-298-695
; Sequence 695, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 695
; LENGTH: 460
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-925-298-695

Query Match          90.6%; Score 87; DB 12; Length 460;
Best Local Similarity 88.9%; Pred. No. 1.6e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 AALEAQCQCIIEYFGDF 18
        |||||:|||||
DB      63 AALEAKICHOIEYFGDF 80

RESULT 6
US-10-102-806-695
; Sequence 695, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 695
; LENGTH: 460
```

GenCore version 5.1.6
Copyright: (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:55:06 ; Search time 40.324 Seconds
(without alignments)
143.151 Million cell updates/sec

Title: US-09-836-073-2

Perfect score: 96
Sequence: 1 AALAEQICQIQEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubppaa/US09C_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubppaa/US10C_NEW_PUB.pep:*
- 16: /cgn2_6/ptodata/1/pubppaa/US10C_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	18	US-09-836-073-2	Sequence 2, Appli
2	87	90.6	18	US-09-836-073-1	Sequence 1, Appli
3	87	90.6	18	US-09-836-073-14	Sequence 14, Appli
4	87	90.6	408	US-10-170-385-477	Sequence 477, App
5	87	90.6	460	US-09-925-298-695	Sequence 695, App
6	87	90.6	460	US-10-102-806-695	Sequence 695, App
7	87	90.6	460	US-10-264-049-2643	Sequence 2643, App
8	83	86.5	17	US-09-836-073-13	Sequence 13, Appli
9	81	84.4	18	US-09-836-073-9	Sequence 9, Appli
10	79	82.3	18	US-09-836-073-11	Sequence 11, Appli
11	79	82.3	18	US-09-836-073-12	Sequence 12, Appli
12	78	81.2	18	US-09-836-073-10	Sequence 10, Appli
13	76.5	79.7	19	US-09-836-073-16	Sequence 16, Appli
14	75	78.1	18	US-09-836-073-3	Sequence 3, Appli
15	75	78.1	18	US-09-836-073-4	Sequence 4, Appli

16	74	77.1	18	US-09-836-073-15	Sequence 15, Appli
17	72	75.0	18	US-09-836-073-7	Sequence 7, Appli
18	71	74.0	18	US-09-836-073-8	Sequence 8, Appli
19	65	67.7	21	US-10-376-121A-20	Sequence 20, Appli
20	63	65.6	18	US-09-836-073-5	Sequence 5, Appli
21	61	63.5	18	US-09-843-676-25	Sequence 25, Appli
22	61	63.5	38	US-09-766-253-25	Sequence 25, Appli
23	61	63.5	38	US-09-438-486-25	Sequence 25, Appli
24	61	63.5	38	US-10-325-810-215	Sequence 25, Appli
25	61	63.5	38	US-10-053-758-25	Sequence 25, Appli
26	61	63.5	38	US-10-054-295-25	Sequence 25, Appli
27	61	63.5	38	US-10-054-611-25	Sequence 25, Appli
28	60	62.5	38	US-10-325-810-214	Sequence 24, Appli
29	58	60.4	16	US-09-836-073-19	Sequence 19, Appli
30	56	58.3	18	US-09-836-073-6	Sequence 19, Appli
31	55	57.3	420	US-10-437-963-128072	Sequence 128072, Appli
32	50.5	52.6	37	US-09-843-676-24	Sequence 24, Appli
33	50.5	52.6	37	US-09-766-253-24	Sequence 24, Appli
34	50.5	52.6	37	US-09-438-486-24	Sequence 24, Appli
35	50.5	52.6	37	US-10-053-758-24	Sequence 24, Appli
36	50.5	52.6	37	US-10-054-295-24	Sequence 24, Appli
37	50.5	52.6	37	US-10-054-611-24	Sequence 24, Appli
38	50	52.1	922	US-10-437-963-125036	Sequence 125036, Appli
39	49	51.0	39	US-09-843-676-26	Sequence 26, Appli
40	49	51.0	39	US-09-766-253-26	Sequence 26, Appli
41	49	51.0	39	US-09-438-486-26	Sequence 26, Appli
42	49	51.0	39	US-10-325-810-216	Sequence 26, Appli
43	49	51.0	39	US-10-053-758-26	Sequence 26, Appli
44	49	51.0	39	US-10-054-295-26	Sequence 26, Appli
45	49	51.0	39	US-10-054-611-26	Sequence 26, Appli

ALIGNMENTS

RESULT 1
US-09-836-073-2
Sequence 2, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Daesupca, Asim
APPLICANT: Daesupca, Asim
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 22002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 18
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match 100.0%; Score 96; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 AALAEQICQIQEYFGDF 18
Db 1 AALAEQICQIQEYFGDF 18

RESULT 2
US-09-836-073-1
Sequence 1, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Daesupca, Asim
APPLICANT: Daesupca, S.

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-24

Query Match 52.6%; Score 50.5; DB 3; Length 37;
Best Local Similarity 83.3%; Pred. No. 0.11; 1; Indels 1; Gaps 1;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICQIEYYFGDF 18
DB 1 ICHQ-EYYFGDF 11

RESULT 15
US-09-430-323-24
Sequence 24, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin
Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-430-323-24

Query Match 52.6%; Score 50.5; DB 4; Length 37;
Best Local Similarity 83.3%; Pred. No. 0.11; 1; Indels 1; Gaps 1;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICQIEYYFGDF 18
DB 1 ICHQ-EYYFGDF 11

Search completed: September 10, 2004, 18:05:07
Job time: 17.8883 secs

FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 214:
US-09-721-456-214

Query Match 62.5%; Score 60; DB 4; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.0029;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICQOIEYFGDF 18
DB 1 ICHQXYFGDF 12

RESULT 13
US-08-851-843A-24
Sequence 24, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809e1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-09-851-843A-24

Query Match 52.6%; Score 50.5; DB 3; Length 37;
Best Local Similarity 83.3%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICQOIEYFGDF 18
DB 1 ICHQ-XYFGDF 11

RESULT 14
US-08-854-050-24
Sequence 24, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6261836e1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536

APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17685
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-214

Query Match 62.5%; Score 60; DB 3; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.0029;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICGQIEYFGDF 18
DB 1 ICHQXEYFGDF 12

RESULT 11
US-09-402-181B-214
Sequence 214, Application US/09402181B
Patent No. 6610839
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 633
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,181B
FILING DATE: 29-Sep-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17685
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ausemus, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 214:
US-09-402-181B-214

Query Match 62.5%; Score 60; DB 4; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.0029;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICGQIEYFGDF 18
DB 1 ICHQXEYFGDF 12

RESULT 12
US-09-721-456-214
Sequence 214, Application US/09721456
Patent No. 6617110
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/721,456
FILING DATE: 22-Nov. 6617110-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017

STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 215:
US-09-402-181B-215

Query Match 63.5%; Score 61; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.002;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICQOIEYFGD 17
Db 1 ICQOIEYFGD 11

RESULT 9
US-09-721-456-215
Sequence 215, Application US/09721456
Patent No. 6617110
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/721,456
FILING DATE: 22-NOV-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 215:
US-09-721-456-215

Query Match 63.5%; Score 61; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.002;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICQOIEYFGD 17
Db 1 ICQOIEYFGD 11

RESULT 10
US-08-974-549A-214
Sequence 214, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:

TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-25

Query Match 63.5%; Score 61; DB 3; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.002;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICQIIEYFGD 17
DB 1 ICEQIEYFGD 11

RESULT 7
US-09-430-323-25
Sequence 25, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin
Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant

MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-430-323-25

Query Match 63.5%; Score 61; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.002;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICQIIEYFGD 17
DB 1 ICEQIEYFGD 11

RESULT 8
US-09-402-181B-215
Sequence 215, Application US/09402181B
Patent No. 6610839
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin B.
Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 633
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,181B
FILING DATE: 29-SEP-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Aueenhue, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid

Query Match 63.5%; Score 61; DB 3; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.002;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICGQIEYFGD 17
Db 1 ICGQIEYFGD 11

RESULT 5
US-08-974-549A-215

; Sequence 215, Application US/08974549A

; Patent No. 6166178

; GENERAL INFORMATION:

; APPLICANT: Cech, Thomas R.

; APPLICANT: Lingner, Joachim

; APPLICANT: Nakamura, Toru

; APPLICANT: Chapman, Karen B.

; APPLICANT: Morin, Gregg B.

; APPLICANT: Harley, Calvin B.

; APPLICANT: Andrews, William H.

; TITLE OF INVENTION: Human telomerase Catalytic Subunit

; NUMBER OF SEQUENCES: 727

; CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A

FILING DATE: 19-NOV-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618

FILING DATE: 01-OCT-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885

FILING DATE: 01-OCT-1997

ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002610US

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 215:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 38 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-974-549A-215

Query Match 63.5%; Score 61; DB 3; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.002;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICGQIEYFGD 17
Db 1 ICGQIEYFGD 11

RESULT 6
US-08-854-050-25

; Sequence 25, Application US/08854050

; Patent No. 6261836

; GENERAL INFORMATION:

; APPLICANT: Cech, Thomas R.

; APPLICANT: Lingner, Joachim

; APPLICANT: Nakamura, Toru

; APPLICANT: Chapman, Karen B.

; APPLICANT: Morin, Gregg B.

; APPLICANT: Harley, Calvin

; APPLICANT: Andrews, William H.

; TITLE OF INVENTION: No. 6261836el Telomerase

; NUMBER OF SEQUENCES: 225

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050

FILING DATE: 09-MAY-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200

PRIOR FILING DATE: 1997-10-06
PRIOR APPLICATION NUMBER: 08/321,427
PRIOR FILING DATE: 1994-10-11
PRIOR APPLICATION NUMBER: 60/086,527
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO: 4
LENGTH: 18
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: B-LAP
US-09-316-630-4

Query Match 90.6%; Score 87; DB 3; Length 18;
Best Local Similarity 88.9%; Pred. No. 4.1e-08;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQICQOIYYFGDF 18
DB 1 AALEAKICHOIYYFGDF 18

RESULT 3
US-08-475-955-20
Sequence 20, Application US/08475955
Patent No. 6641813
GENERAL INFORMATION:
APPLICANT: Harley, John
TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS OF
NUMBER OF SEQUENCES: 218
CORRESPONDENCE ADDRESS:
ADDRESSER: Patrea L. Pabat
STREET: 2800 One Atlantic Center, 1250 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,955
FILING DATE: June 7, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/867,819
FILING DATE: April 13, 1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/648,205
FILING DATE: January 31, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/472,947
FILING DATE: January 31, 1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRP14CIP(2) DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-873-8795
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Binding-site
LOCATION: 8..15
US-08-475-955-20

Query Match 67.7%; Score 65; DB 4; Length 21;
Best Local Similarity 91.7%; Pred. No. 0.00022;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICQOIYYFGDF 18
DB 1 ICQOIYYFGDF 12

RESULT 4
US-08-851-843A-25
Sequence 25, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809e1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-25

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 17:43:01 : Search time 15.8883 Seconds
(without alignments)
58,488 Million cell updates/sec

Title: US-09-836-073-2

Perfect score: 96
Sequence: 1 AALEAKICQIQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	90.6	18	US-09-316-630-3	Sequence 3, Appli
2	87	90.6	18	US-09-316-630-4	Sequence 4, Appli
3	65	67.7	21	US-08-475-955-20	Sequence 20, Appli
4	61	63.5	38	US-08-851-843A-25	Sequence 25, Appli
5	61	63.5	38	US-08-974-549A-215	Sequence 215, Appli
6	61	63.5	38	US-08-854-050-25	Sequence 25, Appli
7	61	63.5	38	US-09-430-323-25	Sequence 25, Appli
8	61	63.5	38	US-09-402-181B-215	Sequence 215, Appli
9	61	63.5	38	US-09-721-456-215	Sequence 215, Appli
10	60	62.5	38	US-08-974-549A-214	Sequence 214, Appli
11	60	62.5	38	US-09-402-181B-214	Sequence 214, Appli
12	60	62.5	38	US-09-721-456-214	Sequence 214, Appli
13	50.5	52.6	37	US-08-851-843A-24	Sequence 24, Appli
14	50.5	52.6	37	US-08-854-050-24	Sequence 24, Appli
15	50.5	52.6	37	US-09-430-323-24	Sequence 24, Appli
16	49	51.0	39	US-08-851-843A-26	Sequence 26, Appli
17	49	51.0	39	US-08-974-549A-216	Sequence 216, Appli
18	49	51.0	39	US-08-854-050-26	Sequence 26, Appli
19	49	51.0	39	US-09-430-323-26	Sequence 26, Appli
20	49	51.0	39	US-09-402-181B-216	Sequence 216, Appli
21	49	51.0	39	US-09-721-456-216	Sequence 216, Appli
22	44	45.8	392	US-08-451-777A-33	Sequence 33, Appli
23	44	45.8	392	US-08-451-777A-33	Sequence 33, Appli
24	44	45.8	392	US-08-998-208-33	Sequence 33, Appli
25	43	44.8	167	US-09-114-000C-3435	Sequence 3435, Appli
26	42	43.8	69	US-09-621-976-7385	Sequence 7385, Appli
27	42	43.8	582	US-09-976-594-733	Sequence 733, Appli

28	41	42.7	29	1	US-08-451-777A-3	Sequence 3, Appli
29	41	42.7	29	2	US-08-451-777A-3	Sequence 3, Appli
30	41	42.7	29	2	US-08-998-208-3	Sequence 3, Appli
31	41	42.7	29	5	PCT-US94-10825-3	Sequence 3, Appli
32	41	42.7	29	5	PCT-US95-06743-3	Sequence 3, Appli
33	40	41.7	141	4	US-09-540-236-2332	Sequence 2332, Appli
34	40	41.7	258	4	US-09-489-039A-11751	Sequence 11751, Appli
35	40	41.7	505	4	US-09-627-216A-12	Sequence 12, Appli
36	40	41.7	505	4	US-09-126-420A-22	Sequence 22, Appli
37	40	41.7	505	4	US-09-765-873A-12	Sequence 12, Appli
38	40	41.7	754	2	US-08-941-262-1	Sequence 1, Appli
39	40	41.7	755	2	US-08-941-262-3	Sequence 3, Appli
40	40	41.7	957	4	US-09-252-991A-21567	Sequence 21567, Appli
41	40	41.7	957	4	US-09-224-024-28	Sequence 28, Appli
42	40	41.7	1180	5	PCT-US94-07902-28	Sequence 28, Appli
43	39	40.6	327	4	US-09-585-858-53	Sequence 53, Appli
44	39	40.6	456	1	US-08-205-719-4	Sequence 4, Appli
45	39	40.6	456	3	US-08-431-517F-6	Sequence 6, Appli

ALIGNMENTS

```

RESULT 1
US-09-316-630-3
Sequence 3, Application US/09316630
Patent No. 6291637
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRRES-MEDIATED TRANSLATION BY A
FILE REFERENCE: 22000-20548-21
CURRENT APPLICATION NUMBER: US/09/316, 630
PRIOR FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: 08/917, 953
PRIOR FILING DATE: 1997-10-06
PRIOR APPLICATION NUMBER: 08/321, 427
PRIOR FILING DATE: 1994-10-11
PRIOR APPLICATION NUMBER: 60/086, 527
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 18
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: LAA
US-09-316-630-3
Query Match 90.6%; Score 87; DB 3; Length 18;
Best local similarity 88.9%; Pred. No. 4.1e-08;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Cy 1 AALEAKICQIQIEYFGDF 18
Db 1 AALEAKICQIQIEYFGDF 18
RESULT 2
US-09-316-630-4
Sequence 4, Application US/09316630
Patent No. 6291637
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRRES-MEDIATED TRANSLATION BY A
FILE REFERENCE: 22000-20548-21
CURRENT APPLICATION NUMBER: US/09/316, 630
PRIOR FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: 08/917, 953

```


This page Blank (uspto)

McGarry, Sean

To: STIC-Biotech/ChemLib
Subject: SEQ SEARCH 09/836,073

Sean McGarry 73484
AU 1635
REM 02D19 Office
REM 2C18 Mailbox
X20761

Please search a. a. SEQ ID NOS: 2-4, 12, and 14-19.

Thank You